3-22-2018

Looking Past the Spark to Find the Fuel of the Arab Spring Fire

Luke M. Brantley

LOOKING PAST THE SPARK TO FIND THE FUEL OF THE ARAB SPRING FIRE

THESIS

Luke M. Brantley, Second Lieutenant, USAF

AFIT-ENS-MS-18-M-111

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

AFIT-ENS-MS-18-M-111

LOOKING PAST THE SPARK TO FIND THE FUEL OF THE ARAB SPRING FIRE

THESIS

Presented to the Faculty
Department of Operational Sciences
Graduate School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
in Partial Fulfillment of the Requirements for the
Degree of Master of Science in Operations Research

Luke M. Brantley, B.S.
Second Lieutenant, USAF

22 March 2018

AFIT-ENS-MS-18-M-111

LOOKING PAST THE SPARK TO FIND THE FUEL OF THE ARAB SPRING FIRE

THESIS

Luke M. Brantley, B.S.
Second Lieutenant, USAF

Committee Membership:

Dr. D. K. Ahner
Chair

Dr. R. R. Hill
Member

AFIT-ENS-MS-18-M-111

Abstract

The field of statistical conflict prediction addresses region-wide analysis in eras of stable conflict and peace. This study improves upon those prediction rates in times of volatile conflict and peace seen during the Arab Spring of 2011 to 2015. During this time, higher rates of conflict transition in certain Middle Eastern and North African countries occurred than normally observed in previous studies. Due to the fact that previous prediction models decrease in accuracy during times of volatile conflict transition and since the proper strategy for handling the Arab Spring has been highly debated, this study considers alterations to previous studies to understand the effects of the Arab Spring on conflict prediction over a five-year period. This study identifies which countries were affected by the Arab Spring, and then applies logistic regression to predict a country's tendency to suffer from high-intensity, violent conflict. A large number of open-source variables are incorporated by implementing an imputation methodology useful to conflict prediction studies in the future. The imputed variables are implemented in four model building techniques: Purposeful Selection of Covariates, Logical Selection of Covariates, Principal Component Regression, and Representative Principal Component Regression resulting in accuracies exceeding 90%. Analysis of the models produced by the four techniques supports hypotheses which propose political opportunity and quality of life factors as causations for increased instability following the Arab Spring.

AFIT-ENS-MS-18-M-111

I dedicate this work to my friends and family for their love and support.

Acknowledgements

I would like to express my sincere gratefulness to my advisor, Dr. Darryl Ahner, for his expertise, insight, and mentorship in my thesis process.

I would also like to thank my classmates for their ability to calm my fears and direct me in the right path throughout my Master's Degree process.

Luke M. Brantley

Contents

List of Figures

List of Tables

Table Page

LOOKING PAST THE SPARK TO FIND THE FUEL OF THE ARAB SPRING FIRE

I. Introduction

1.1 General Issue

From the self-immolation of Muhammad Al-Bouazizi to the prolonged occupation of the Islamic State of Iraq and Syria (ISIS), what is now known as the Arab Spring caused a complex shift in stability in a subsection of the Middle East and North African region. This sudden shift which saw varied versions of regime change for four nations in the region, was unexpected by the US and the rest of the western world [2]. The vacuums of power seen in these regime changes were filled with chaos in some nations, and stability in others. Other countries still responded to the atrocities in the region and called for reformation without toppling regimes and saw similar, varied results. The reactionary policies by the rest of the world, enacted following the initial sparks of instability in 2011, have been highly debated. This paper hopes to aid decision makers with postdictive analysis on how to most effectively allocate resources and alter policy outcomes in the latent environment of future high-conflict regions. It has already been shown by Boekestein [3], Shallcross [4], and Leiby [5] that conflict can be predicted using Political, Military, Economic, Social, Infrastructural, and Information Systems (PMESII) data. This study furthers their research by focusing on a specific anomaly of conflict shock to a region known as the Arab Spring.

1.2 Problem Statement

Using open-source, country data, is it possible to better understand why countries, spurred by an Arab Spring-like event, that see a region-wide increase in protests, remain or enter into "high-intensity, violent conflict" over five years or settle into lower levels of conflict [1]?

1.3 Research Questions

To answer the problem statement, this paper addresses the following statistical and political research questions.

Statistical

S1. What is the best method for imputing missing PMESII data, even when a large portion of the most recent data is missing?

S2. What is the most effective method for model building to capture the tendency of a country to fall into conflict?

Political

P1. How can nations be identified as being affected by the Arab Spring?

P2. What open source PMESII factors can be identified that affect the selected Arab Spring nations' tendencies to transition into and out of conflict?

P3. Can nations receiving a conflict shock be grouped into two groups: at-risk for escalated conflict and not at-risk for escalated conflict over a five-year period following the shock?

P4. Could the probability of long-term, escalated conflict have been decreased following the onset of the Arab Spring by altering certain PMESII factors?

1.4 Methodology Overview

This study combines data from multiple open-source data sets to allow for an accessible answer to the research questions. Partially due to the differences in the combined data sources, multiple imputation was used to fill in the missing data in the combined database. Using the completed data, trend variables and new variables were created from the original data through Principle Component Analysis (PCA). All of these variables were then used as indicators to build a logistic regression model against a binary conflict variable. The model was built using four different methods: Purposeful Selection of Covariates (PSOC), a variant of PSOC which is named Logical Selection of Covariates (LSOC), Principal Component Regression (PCR) and a derivation of PCR which is named Representative Principal Component Regression (RPCR). The model was validated through Pearson χ^2 tests, Hosmer-Lemeshow tests, Receiver Operator Characteristic (ROC) Curves, and classification table analysis. With the validated model, sensitivity analysis helps to determine how policy makers could make helpful changes to the conflict environment if another situation like the Arab Spring arose. Figure 1 depicts the direction of the study which illustrates the three general efforts made: data preparation, modeling, and analysis.

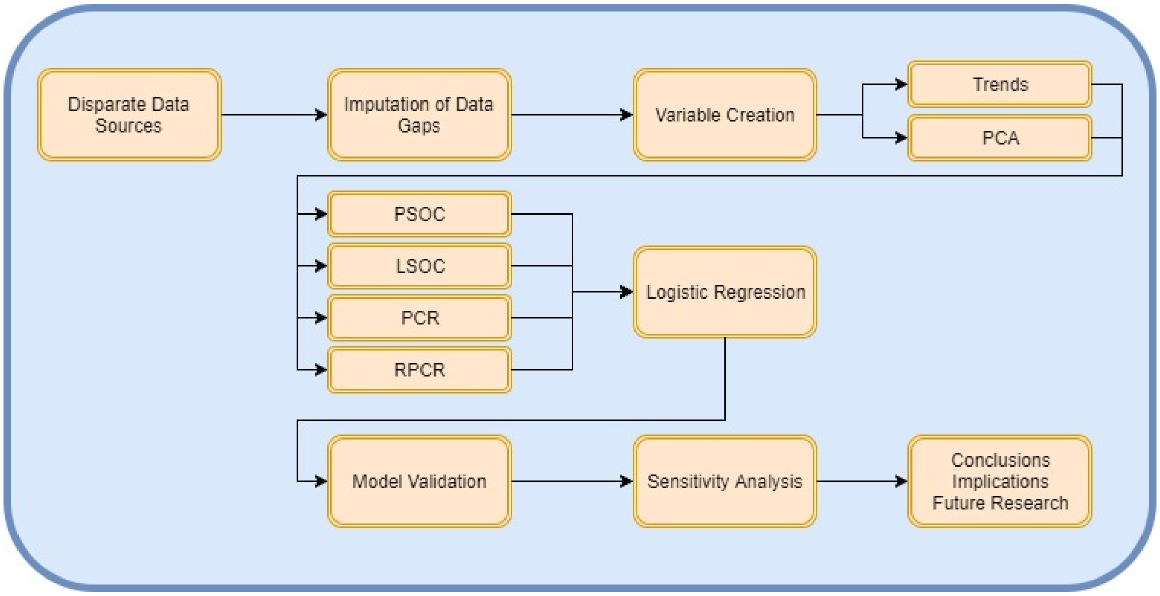

Figure 1. Flow Chart of Methodology

1.5 Implications

This study should provide decision makers a useful model to predict which countries are at risk of violent conflict for the next five years following an initial conflict shock. It should also provide decision makers insight into the pertinent tools for adjusting these probabilities. For future researchers, this study provides meaningful methodology for imputing recent data and data missing due to political turmoil, as well as a model building strategy for the complex PMESII environment.

1.6 Paper Outline

The rest of this paper is organized into four additional chapters. Chapter 2 provides an overview of relevant prior research of the historic environment surrounding the Arab Spring, statistical methods used throughout our study, as well as previous conflict prediction studies. Chapter 3 walks the reader through the methodology used from data cleaning to logistic regression model building. Chapter 4 follows to validate the methodology and conduct sensitivity analysis on the resulting logistic regression model to address the problem statement. The paper is finished in Chapter 5 with a conclusion of the methodology, analysis, and results.

II. Literature Review

2.1 Overview

This paper addresses findings pertinent to both statistical and political science fields. Because of this range of audience, the literature review attempts to bridge the gap between explaining the historical and social environment which is now called the Arab Spring and the predictive methods which will be used in order to gain insights from our methodology. To do this, this section paints the picture of the events leading up to and occurring in the 2011 Arab Spring, then explains the specific statistical methodology used in the analysis, and finally cites previous works that meld these two fields together showing the ability to predict conflict levels.

2.2 Political Science

Defining the Arab Spring Region

This study is specifically interested in determining what caused the differences in outcomes between the countries still facing conflict stemming from the Arab Spring and those that now experience lower levels of conflict. This effort required defining the set of countries to analyze in order to capture all countries that were affected by the Arab Spring even if they are not currently feeling the effects from it now. In essence, this defines the Arab Spring. The study, “Bread, Justice, or Opportunity? The Determinants of the Arab Awakening Protests”, studies the up-tick in protests from 2006 to 2011 in the Middle East and North Africa [6]. It identifies countries that experienced an increased number of protests calling for government reform in its definition of the “Arab Awakening”. Their study shows this occurred in eleven countries: Bahrain, Egypt, Jordan, Kuwait, Libya, Mauritania, Morocco, Oman, Syria, Tunisia, and Yemen.

Table 1. Arab Spring Countries

Bahrain	Egypt
Jordan	Kuwait
Libya	Mauritania
Morocco	Oman
Saudi Arabia	Syria
Tunisia	Yemen

This work also adopts this list along with Saudi Arabia as the countries considered to be affected by the Arab Spring as these countries experienced an increase in political demonstrations, both violent and non-violent, following the toppling of Tunisia's President in January 2011 [6]. Through analysis of the events of the Arab Spring, Saudi Arabia is included due to their government's harsh reaction to initial protests which diminished quickly although not from a lack of fervor from protesters. This signifies Saudi Arabia as an interesting case in the diverse reaction to the Arab Spring. In general though, an increase in demonstrations is a useful standard for consideration as it is recognizable for future application and describes the tools by which change occurred following the Arab Spring. This answers research question P1, how can nations be identified as being affected by the Arab Spring?

Overview of the Arab Spring

To fully understand how conflict can be predicted given the initial state of protest in the Arab Spring, a thorough understanding of the events that transpired leading to its spark must be observed. First, the historical facts. Chapter 2 of The Future of the Arab Spring uncovers some of these historical factors leading up to the Arab Spring from the 1980s when affected nations began having problems dealing with volatile oil prices [7]. It also describes more immediate factors leading to the Arab Spring where the growing young, highly educated population grew in frustration as unemployment reached worldwide highs. This study seeks to understand the characteristics that tended nations into or away from conflict following the onset of the Arab Spring so we find it pertinent to provide an overview of the events in each country during this period following the spark as well.

The spark of the Arab Spring began with the self-conflagration of Mohamed Bouazizi, a street vendor in Tunisia, who was frustrated and humiliated by a female government official. His passion was indicative of the rest of the country's frustration as they responded with protests in multiple parts of the country including the capitol, Tunis. Within a month, amidst continued protests, the Tunisian president of twenty-three years fled the country. Perhaps due to the perceived success of the Tunisian protests and due to frustrations with their own government, thousands of Egyptians protested during their National Police Day just over a month after Bouazizi's martyrdom. Violent responses to Egyptian protests provided enough motivation for the population to overthrow their president of thirty-one years within the next month. The two successful regime changes in Tunisia and Egypt signaled to the rest of the region that factions had an opportunity to address the problems they perceived in their own countries. Weyland [8] argues though that this success gave citizens in nearby countries a false sense of hope that they too could make a similar type of change just as easily in those countries, however they did not take into account the factors that this study attempts to uncover that tend a country toward further destabilization rather than peace when the government is tested.

Many countries saw some sort of change during the Arab Spring, ranging from new policies to new regimes as a result of both peaceful protests to full scale civil wars. Below is a short description of the outcomes of the Arab Spring for all affected countries.

The King of Jordan was able to halt calls for him to step down as King by sacking the existing parliament under him [9]. Jordanians calmed their protests after this as their larger middle class did not desire to tear down their monarchy and were pleased with their economic conditions.

Libya faced a long, violent campaign to overthrow their leader, Muammar Gaddafi, and has since observed massive struggles from competing governing bodies and ISIS [10].

The Arab Spring also had a significant presence in Bahrain, however the Shiite protesters were quickly put down as Saudi troops helped the Sunni monarchy in Bahrain violently quell any possible uprisings [11]. The Saudi Arabian government feared that the protests in Bahrain would spur a similar movement from the Shiite minority in their own country and would give Shiite-controlled Iran more power in the region.

These protests in Bahrain were closely related to the protests seen within Saudi Arabia which were likewise extinguished with a quick and overpowering majority opposition from the Saudi government [12].

Kuwait had a similar, inconsequential revolution attempt as only a minority of the population desired government reformation and protests were met with police opposition and imprisonment for some [13]. Like Jordan, the Emir of Kuwait changed the voting process and saw the country through a change in prime minister as well as many other top government positions, but maintained most of the same governing characteristics.

The Arab Spring in Morocco was also similar to that of Jordan's with the King giving some power to citizens in response to large protest turn-outs, but not fully relinquishing control nor acquiescing to all demands [14].

Protests in Mauritania seemed to be quite related to those in Tunisia as they were sparked by another self-immolation of a Mauritanian man who provided a list of reforms he requested from the president [15]. As with other countries in the region, protests followed, however, all of them peaceful and requesting changes in policy instead of changes in leadership. Their pleas for increased human rights efforts were not met with strong change, however, in-country awareness and desire for change increased among the population.

In Oman, protests were spurred more specifically due to economic issues in northern Oman for the region's youth [16]. These protests, although initially met with violence from the government did eventually bring limited change. The Sultan increased the minimum wage, created 50,000 new jobs, and replaced ten ministers in the Sultan's cabinet.

Syria's history following its own version of reformation-hopeful protests has been plagued by violence and civil war from multiple sides with multiple foreign influences [17]. The same reactionary disgust spread throughout Syria as protests were met with violent police brutality. Opposition rose against President Assad who used military assistance to put down organized protests posing a significant threat to the Assad regime. The Syrian situation differed however as foreign powers such as Russia and Iran backed the Assad regime while the US and Great Britain recognized opposition regimes creating tension which allowed a civil war with large civilian casualties to sweep the country. The other nuanced part of Syria's Arab Spring was the rise of ISIS in 2014. The competing parties of ISIS, Assad, and opposition groups has flung Syria into prolonged, high-intensity conflict since the beginning of the Arab Spring.

Yemen is another country which has experienced high-intensity, violent conflict following similar hopeful protests at the beginning of the Arab Spring [18]. Initially, the Yemeni revolution was successful with the ouster of President Ali Abdullah Saleh who had been in power for 33 years, however Saleh came back after the election of a new president and allied with the Houthis, a Shiite religious sect, to overthrow the capital Sana'a. Since then, violence has proliferated in the country between Houthis, along with Saleh's supporters, and the recognized government of Yemen backed by foreign powers such as Saudi Arabia.

Hypothesis Building Research

As its title suggests, "Bread, Justice, or Opportunity? The Determinants of the Arab Awakening Protests" by Costello et al. focuses on finding the causes of the explosion in the number of protests which dominated the early days of the Arab Spring [6]. Using news reports as an indicator for protest frequency, it tests four commonly debated hypotheses regarding the causes of the Arab Spring. These four hypothesized factors are a growing, young, educated population, "democratic deficit" or long-standing authoritarian rule, political opportunity, and "the growth of the new communications media" through cell phones and the Internet. This paper also considers these factors in the model building process.

Costello et al. [6] argue that the growing amount of young, unemployed citizens in the region spurred economic incentives to protest against standing governments. Because of this, they included a variable capturing the percent population between the ages 15 to 29, but state that a better indicator would be the unemployment rate of this age group. This argument is supported by Urdal [19] who completed a full analysis of six hypotheses of factors that further explain the conflict creating nature of the youth bulge or a large population of 15 to 29-year-olds in a country. He cites previous theoretical studies behind phenomenon and concludes that a country's tendency towards armed conflict increases in the presence of a youth bulge when the age dependency ratio, or percent of the population under the age of 15, is low or when the country is either highly autocratic or highly democratic. He also found that countries with a youth bulge are more likely to experience increased terrorism when coupled with either low economic growth or high tertiary (college) education participation. Both of these factors lead to higher unemployment or at least, less jobs matching the large youth population's skill sets which, he suggests, decreases the opportunity cost to obtain alternative incomes and hope through rebel or terror groups. The final hypothesis Urdal tests, whether urbanization increases or decreases likelihood of conflict, was found to be insignificant for both armed conflict as well as terror. Costello et al. were not able to follow Urdal's prescription of using the youth cohort's unemployment rate instead of percent population to test the youth bulge effect and consequently found the youth bulge to be insignificant in the case of the Arab Spring. We find this result unlikely for at least sparking conflict. Because we are able to include unemployment rates of the youth cohort, we continue to analyze this factor in our modeling.

The conclusions by Costello et al. [6] on economic and political factors interestingly found economic factors more as a channel in which politically fueled grievances were voiced. The most significant factors in the rise in protests during this time seem to be caused by political terror and the opportunity to voice political dissension from previous gains in civil liberties and the decentralization of the media. This identifies that a significant predictor for similar uprisings is the growth of political opportunity or the reduction of barriers for the masses to voice their grievances. The gradual nature of these identifying occurrences lends credibility to the unexpectedness of the Arab Spring.

Costello et al. [6] also assert that the use of cell phones and social media is an important consideration for the escalation of protests. Wolfsfeld et al. [20] challenges the propensity to credit the advent of social media as a main impetus for the increase in severity and violence seen in the Arab Spring. Their study hypothesizes that widespread media is a reaction and not a cause of political uprising and that the effect of social media is different depending on the political environment in which the protests occur. Khondker [21] focuses on this idea and analyzes the effect of technology and social media as an impetus for the efficacy of the Arab Spring. This study found an interesting role that technology does have some efficacy as an instrument for change but that governments can also use these tools to further repress the masses. Because of this contention, technology use is analyzed in this study.

Anecdotal Hypothesis of Conflict Escalation Causation

Although the factors reviewed in the previous section may be sufficient for starting conflict, this study is interested in predicting prolonged conflict which requires both spark and fuel. Part of the fuel that continued so many of these conflicts came from the occupation of terror groups, namely ISIS, or other rebel groups as explained in our Arab Spring profiles as well as by Urdal's explanation for how poor economic growth and high tertiary education create more terrorism [19]. This section suggests a hypothesis for how an initial shock of conflict in a region can lead to prolonged, high-intensity, violent conflict. Stemming from the political science research above, this study posits that there are three main factors that cause protests to escalate to civil war. These factors are the severity of the protesters' desires, the willingness of the current regime to concede to the protesters' desires, and the potential for a power vacuum.

The first of these, the severity of protesters' desires can be contrasted between Saudi Arabia and Tunisia. In Saudi Arabia the Shiite minority that protested were calling for fairer treatment of their own group and therefore did not garner a large following from within the country. In Tunisia thousands of protesters poured into the streets in front of government buildings throughout the country with calls for reform quickly turning into calls for revolution following the government's violent response

[22]. There were also differences in motivations for protests. Some called for overhauls of constitutions and leaderships while others simply prodded for larger emphasis on economic issues.

The second hypothesized factor, the willingness of a regime to concede to protesters, stems from two sub factors, corruption level, and level of autocracy. Corruption level is often expressed by the number of years the regime has been in office. Autocracy level explains the tendency of a nation towards war based on the presence of scapegoats for current problems and the extent to which the regime responds to protests with violence. Several of the overthrown regimes led their countries for over 30 years and military alliances with these leaders ran deep. Some, like the King of Jordan, responded to demands by sacking the senior government leaders under them while others were unable to shift blame to others especially when ordering police and military to fire upon protesters as seen in many of the Arab Spring countries. This factor is also similar to what Costello et al. [6] considered political opportunity. The first two factors have the largest correlation with toppling the current regime as that was usually the impetus for further escalation as explained by the third factor.

The third factor, the potential for a political vacuum is the greatest determining factor for prolonged civil war. It is determined by two sub factors, the availability of a peaceful transition of power and the reasons for reform. The first of these is often unknown at the time of transition but can depend on the effectiveness of the incoming leader, the presence of competing parties for new leaders, and the methods in which the new leader assumes power. The second sub-factor is a causal influence on prolonged conflict. Protests which only called for policy reform often were settled whereas ideological challenges more often escalated and were supported by violent groups such as the Houthis in Yemen and ISIS in Syria. Following a continued review of the statistical tools used in our research and relevant prior works in this conflict prediction, we test these hypothesized factors.

2.3 Further Development of Necessary Statistical Tools

Logistic Regression Overview

Various mathematical techniques are used in this study centered around logistic regression with a binary dependent variable. The book *Applied Logistic Regression*, by Hosmer, Lemeshow, and Sturdivant [23] provides an in-depth understanding of what logistic regression is and how to apply it to real data. With dichotomous responses, a linear regression would not provide accurate understanding of the variance in observations and would suggest predicted values outside of the possible set [0,1]. Instead, logistic regression uses the model shown in Equation 1 to determine the outcome, the expected value of the dependent variable conditioned on the independent variables.

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} \tag{1}$$

This model form yields an equation for the outcome, Y given by $Y = \pi(x) + \epsilon$ where ϵ is the error which follows a binomial distribution with mean zero and constant variance $\pi(x)(1-\pi(x))$ differing from the normally distributed error of linear regression. This model satisfies the assumption that outcomes must be greater than or equal to zero and less than or equal to one. An important transformation of the logistic regression equation reveals the power of the model. The logit transformation, seen in Equation 2, transforms the nonlinear equation into the linear regression equation.

$$g(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_1 x \tag{2}$$

The estimates of $\pi(x)$ are found using maximum likelihood estimation instead of least squares estimation since $\pi(x)$ is a nonlinear function. Maximum likelihood estimation is the method which produces least squares estimation for linear models and for logistic regression, maximizes the probability that the function created from the estimated parameters matches the observed data. These techniques drop several of the assumptions seen in linear regression. These dropped assumptions include

- Errors have mean of zero
- Linear relationship between dependent and independent variables
- Constant Variance (Homoscedasticity)
- Normal errors

Instead, logistic regression requires the following assumptions [24]:

- Independent errors
- Binomial errors
- Linear relationship between independent variables and the logit function
- Correct model specification
- Limited multicollinearity
- Larger number of observations required for convergence (5-10 observations per independent variable)

The estimated parameters represent the amount that each independent variable increases the probability that $y = 1$. In this study these estimates indicate the change in probability that an increase of an independent variable has on a country being in conflict.

Model Building Strategies

While considering the intricacies of logistic regression and our problem, one must also understand the benefits and disadvantages of different model building strategies. To provide a full analytical understanding of the current problem, four building strategies, Purposeful Selection of Covariates, Logical Selection of Covariates, Principle Component Regression, and Representative Principle Component Regression are employed.

The first model building method was Purposeful Selection of Covariates (PSOC) as defined by Hosmer, Lemeshow, and Sturdivant [23]. PSOC is a seven-step process that alternates between adding and removing indicators with a focus on model significance, coefficient significance, and correlation between variables. It is superior to stepwise regression as it requires judgment from the analyst to distinguish between the importance of variable significance and possible collinearity. The process of adding and removing variables also helps ensure that variables that help refine other variables are included and not only those with large effects on the dependent variable. The first step selects all indicators, whose univariate models are significant at a high (≥ 0.25) α level. Second, a multivariate model is made from the previously selected indicators and is reduced until all coefficient estimates are significant at the 0.05 α level. The now reduced model is compared to the multivariate model from the beginning of the second step using a partial likelihood ratio test. This likelihood ratio test compares the reduced model to the full model according to Equation 3.

$$D = -2\ln\left[\frac{likelihood\ of\ the\ reduced\ model}{likelihood\ of\ the\ full\ model}\right] \tag{3}$$

Where the likelihood function is defined by Equation 4 with $\pi(x_i)$ being the predicted probability for observation i and y_i being the actual dependent variable outcome for observation i.

$$likelihood = \prod_{i=1}^{n} \pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i} \tag{4}$$

The D statistic is compared to a χ^2 distribution with the degrees of freedom equal to the difference in the number of variables between the full and reduced model. If the test of the deviances of these two models is significant (< 0.05), then the reduced model is not as accurate within statistical significance so indicators should be added back to the reduced model from the full model. If the test is insignificant (> 0.05), then the reduced model is just as good of a predictor as the full model and should be kept to retain parsimony. In the third step, the magnitude of the coefficients before and after reducing the model is checked. If these magnitudes change by more than 20-25%, then indicators should be added back in from the full model to determine if this can be rectified or to understand how collinearity is affecting the estimates.

The fourth step involves reinserting indicators deemed insignificant in the univariate tests. This is necessary as some indicators may exhibit confounding effects with other main effects. In the fifth step, the linearity of the indicators compared to the logit function are checked. There are several methods to check the logit linearity assumption proposed in *Applied Logistic Regression* [23]. However, in most cases it is sufficient to observe possible linearity in plots of the independent variables to the predicted values from logistic regression. The sixth step checks for significant interactions of variables already in the model at the end of step five. Finally, models using PSOC must be validated in step seven.

The first model follows the indications of PSOC based solely on the statistical findings of each step without further analytic intervention. PSOC was extended into logical selection of covariates (LSOC) as each step of PSOC is interrupted to understand how the current model supports or opposes the anecdotal hypothesis from research. This is initialized by only considering variables that relate to one of nine categories that define our anecdotal hypothesis or were outcomes from previous conflict prediction studies. The nine categories relevant to the research hypotheses are displayed in Table 2.

Table 2. Variable Categories for LSOC Consideration

Category
Autocracy
Corruption
Opportunity
Quality of Life
Regional Effect
Technology
Threat of Protest
Violence of Response
Youth Bulge

While the benefit of the pure PSOC method is that it allows for variables not captured in our hypothesis to be tested, the benefit of LSOC is that its outcome should provide more interpretable results and can be compared to previous studies.

The last two models, PCR and RPCR rely on the computation of principal component analysis (PCA). PCA is a data reduction technique that creates new variables, called components, from linear combinations of the input variables. PCA is used in this case to represent the multiple correlated variables in the dataset with uncorrelated components. The relationship between input variables Y_i, and principle components C is shown in Figure 2 where w_i are the weights used to calculate the linear combination of input variables to create the principal components [25].

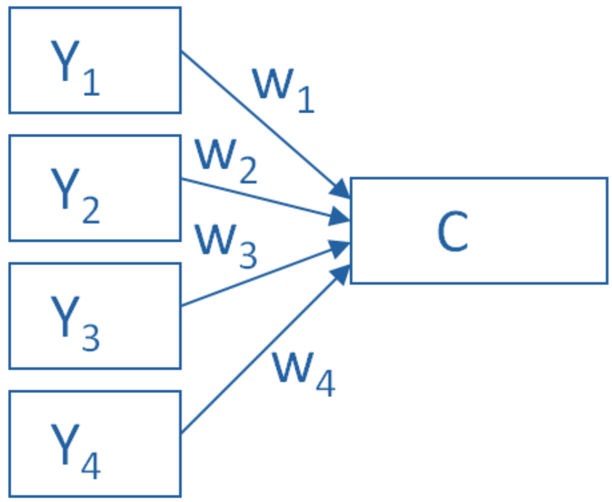

Figure 2. Relationship of Variables to Principle Components [25]

These linear combinations are formed so that each component is orthogonal to all previous components. This is done iteratively by first finding the linear combination with the largest eigenvalue and then repeating this process for all successive components. Since our data have varying ranges for different variables, we first standardize all variables using Equation 5.

$$Zscore = \frac{x - \overline{\mu}}{\sigma} \tag{5}$$

The result is p independent variable orthogonal components. When using PCA for dimensionality reduction, a subset of the first $m < p$ should be used to represent the data. However in our discussion of principle component regression, we make the claim that we must consider all components until there is further understanding of the components.

Principle Component Regression (PCR) uses the components from PCA as the independent variables for regression. The greatest advantage of PCR is that it groups collinear variables into orthogonal components so multicollinearity is no longer a problem. It also incorporates all variables in the dataset when research may have missed significant variables, like PSOC. As alluded to in the previous paragraph, all components must be analyzed for entry into the model, not just those that account for the largest percent of variance in the independent variables. According to Had, Ling [26] it is possible for the principle component (PC) with the smallest eigenvalue to contain most or all of the explanatory power of the dependent variable in PCR . They explain this occurs when β is in the direction of the eigenvector corresponding to this PC. To overcome this shortfall of PCR, PCs were included in PCR according to a sufficient univariate fit with the dependent variable as in the first step of PSOC. This led to inclusion of PCs with small eigenvalues that exhibit better fit to the dependent variable.

To improve the interpretability of the model with respect to PCR, a new model building technique expanding upon the idea of PCR was created. Representative Principle Component Regression (RPCR) takes the relevant PCs, determined by univariate testing, and chooses one representative variable, that explains the most variance, from each PC and uses this variable as the sole representative of the orthogonal PC to consider in the model building process. Interactions between representative variables are also included where clear differentiation between in-conflict and not-in-conflict observations can be determined through the use of PCA score plots. This should incorporate the benefit of multicollinearity reduction from PCR and the benefit of model interpretability from covariate selection methods.

The Wald test is used to test the significance of all j variables. This test takes the square of the Wald Statistic in Equation 6 and compares it to the χ^2 distribution with one degree of freedom [23]. Under the null hypothesis of this test, the variable has no significant contribution to the model so a small p-value (< 0.05) indicates that the variable does provide a significant contribution to the model.

$$W = \frac{\hat{\beta}_j}{\widehat{SE}(\hat{\beta}_j)} \tag{6}$$

All models are checked for adequacy using Pearson Chi-squared tests, Hosmer Lemeshow tests, and receiver operating characteristic (ROC) curves as well as other sensitivity analysis driven validation techniques. In these tests, the assumption is that the model is correctly specified in an attempt to determine if the predicted probabilities accurately reflect the true outcome experienced in the data. The Pearson Chi-squared test uses the idea of the residual, or distance between predicted and observed values. The equation for the Pearson residual is seen in Equation 7.

$$r(y_j, \hat{\pi}_j) = \frac{y_j - m\hat{\pi}_j}{\sqrt{m\hat{\pi}_j(1-\hat{\pi}_j)}} \tag{7}$$

Where m_j is the jth unique observation and π_j is the prediction from the logistic regression equation. The test then determines if the calculated residual follows the assumed Chi-squared distribution with $J-(p+1)$ degrees of freedom with J equal to the number of unique observations and p equal to the number of coefficient estimates using the statistic in Equation 8.

$$\chi^2 = \sum_{j=1}^{J} [r(y_j, \hat{\pi}_j)]^2 \tag{8}$$

The Pearson Chi-squared test has two shortcomings in this study. First, it provides a general assessment of the model fit over the entire range of predicted values and may not pick up poor model fit for one region of predicted values if another region had superior model fit. Second, the p-values obtained from the Chi-squared test are not reliable when $J \approx n$, with n equal to the total number of observations, which is the case in this study as no two country-year observations are the same for almost every combination independent variables.

The Hosmer-Lemeshow (H-L) test avoids both issues as it follows the Chi-squared distribution with $g - 2$ degrees of freedom, with g = number of groups in the calculation, and by applying the Pearson Chi-squared test to multiple regions of the range of predicted values. The number of groups, g, in this study is based on ten percentile cutoffs or "deciles of risk" as suggested by Hosmer, Lemeshow, and Sturdivant [23]. The H-L test relies on applying the Hosmer-Lemeshow goodness of fit statistic in Equation 9a, with c_k being the number of unique observations in the kth group, which compares the expected and observed number of observations within each group.

$$\hat{C} = \sum_{k=1}^{g} \left[\frac{(o_(1k) - \hat{e}_1 k)^2}{\hat{e}_1 k} + \frac{(o_(0k) - \hat{e}_0 k)^2}{\hat{e}_0 k} \right] = \sum_{k=1}^{g} \left[\frac{o_(1k) - n_k \overline{\pi}_j}{n_k \overline{\pi}_j (i - \overline{\pi}_j)} \right] \tag{9a}$$

$$o_(1k) = \sum_{j=1}^{c_k} [y_j] \tag{9b}$$

$$o_(0k) = \sum_{j=1}^{c_k} [m_j - y_j] \tag{9c}$$

$$e_(1k) = \sum_{j=1}^{c_k} [m_j \hat{\pi}_j] \tag{9d}$$

$$e_(0k) = \sum_{j=1}^{c_k} [m_j (1 - \hat{\pi}_j)] \tag{9e}$$

$$\overline{\pi}_j = \frac{1}{n_k} \sum_{j=1}^{c_k} [m_j \hat{\pi}_j] \tag{9f}$$

A high p-value from the Chi-squared test suggests that the model fits well however p-values close to, but above 0.05 should be looked at with caution for conclusion of model fit. The findings of the H-L test should also be questioned if the number of estimated values in any groups is less than five as conclusions of these groups' adequacy may not be well founded.

The final statistical concept reviewed in this section is that of ROC curves. To understand the results from ROC curves, one must first understand the concepts of sensitivity and specificity. Originating from classification table analysis, these terms relate the model's ability to correctly predict observations as either 1 or 0 in logistic regression based upon a given cutoff of the predicted values. Sensitivity provides the ratio of correctly predicted occurrences of an event (1) to the total number of occurrences of an event. This can be thought of as "the probability of detecting a true signal." Specificity provides the ratio of correctly predicted non-occurrences of an event (0) to the total number of non-occurrences of an event. This can be thought of as "the probability of detecting a false signal." The issue with using sensitivity and specificity as pure validation techniques is that the cutoff for determining a true signal and a false signal is arbitrary and will most likely change the outcome of both of these measures as the cutoff is changed. ROC curves overcome this pitfall by observing these measures over the entire range of possible cutoff values $[0, 1]$. The ROC curve is formed by plotting sensitivity versus (1-specificity) for cutoffs 0 through 1. The area under the curve (AUC) is then calculated and used to determine model adequacy. A well excepted table of guidelines to interpreting the AUC is given in Table 3. This guide can be found in [23].

Table 3. AUC Discrimination Levels

AUC = 0.5	No Discrimination
0.5 < AUC < 0.7	Poor Discrimination
$0.7 \leq AUC < 0.8$	Acceptable Discrimination
$0.8 \leq AUC < 0.9$	Excellent Discrimination
$AUC \geq 0.9$	Outstanding Discrimination

2.4 Applying Statistical Methods to Conflict Prediction

Our study expands upon the work of Boekestein [3], Shallcross [4], and Leiby [5] in the area of conflict prediction. These studies introduce many of the important features and methods used in conflict prediction and provides a baseline for prediction accuracy. This study applies similar methodologies in a unique application of conflict in the Arab Spring.

When analyzing international conflict, one widely accepted standard of conflict definition comes from the Heidelberg Institute for International Conflict (HIIK). The HIIK rates the highest level of conflict that each country has experienced in a year on a 0-5 scale. Each level is identified by the following set of terms: no conflict, dispute, non-violent crisis, violent crisis, limited war, and war. Shearer and Marvin [27] were the first to condense this conflict barometer rating into a binary variable of in-conflict and not-in-conflict for use in a logistic regression. Boekestein [3] borrows this dependent variable classifying conflict with the last three levels, violent crisis, limited war, and war, and not-in-conflict as the bottom three classifications, no conflict, dispute, and non-violent crisis, whereas Shallcross [4] and Leiby [5] use a conditional conflict transition dependent variable. These last two studies concerned themselves with predicting how countries transition into or out of a state of conflict as opposed to predictions of general status of conflict. This study focuses on countries attaining pure conflict levels and is less concerned with the transitional properties of conflict so it adopted a variation of the Boekestein [3] and Shearer, Marvin [27] dependent variable as discussed in 3.6.

The open-source variables used in this study, outlined in Section 3.7, are gathered, stored, and maintained by AFIT. Many of the candidate, independent variables are shared with previous AFIT studies, however other variables used in these theses could not be included in this study due to data limitations. All three of these studies performed a global perspective of conflict incorporating 180 to 182 countries' data. Since this study only observes twelve nations, imputation was not possible for variables such as Improved Water Availability as some countries had no current or historical records of these variables. The candidate variables that were considered in this study come from several different open source data bases except for Percent Border Conflict and Average Border Conflict which were derived by Boekestein [3]. These variables attempt to capture the effect that neighboring countries' conflict levels have on a country. On a tactical level this was an observably significant factor in the Arab Spring in sparking protests in nearby nations as several protesters referenced the previous actions of other nations' protests in their cries for change [28].

Previous AFIT theses were used as a baseline to judge the impact of new prediction rates although the prior methodologies and data sets are not identical to those of this study. In Boekestein's model for "Arab Countries" he obtained an 84.31% prediction accuracy on the training set and 70.59% prediction accuracy on the validation set. The training set consisted of 180 countries from 2011 to 2012 and the validation set consisted of the same countries for 2013. This model consisted of five variables, Death Rate, Arable Land, Refugees (Asylum), Trade, and Freedom. Shallcross took the logistic regression methodology and implemented the probabilities from the model to inform a Markov Chain framework to predict conflict transition. His methodology for the "Arab & North African States" region produced a 93.72% prediction accuracy on the training set and 70.69% prediction accuracy on the validation set. This validation set accuracy for the Arab and North African States was much lower than all other regions' validation accuracies which Shallcross notes is most likely due to the abnormality of the Arab Spring. The training set consisted of 182 from 2004 to 2010, and the validation set included the same countries from 2011 to 2013. This is a problematic data split as it separates the pre-Arab Spring environment from the post-Arab Spring. The Shallcross logistic regression model for the region closest to the current model consisted of four variables, Ethnic Diversity, Regime Type (Democratic), 3 Yr Freedom Trend, and Regime Type (Emerging). The Leiby study follows a similar process to Shallcross using a conditional conflict transition dependent variable but does not incorporate Markov Chain modeling, and instead looks at forcing border conflict and Fresh Water per Capita into the model. The maximum prediction accuracy that his models obtain is 96.15% for the in-conflict training set and 93.02% for the in-conflict validation set, however for the not in-conflict validation set his model obtains 66.67% accuracy. The Leiby training and validation sets were the same as those used by Shallcross. These models included five variables, Birth Rate, Death Rate, Government Type (Autocratic), 3 Yr Freedom Trend, and Ethnic Diversity. Since our model captures only the anomaly era and region of the Arab Spring and previous theses capture a larger scope, we hope to improve upon these accuracy statistics.

2.5 Summary

This literature review bridged the gap for political science and statistical audiences. This involved first describing some historical facts and the leading hypotheses surrounding the severity of the Arab Spring. It then provides basic approaches of logistic regression and model building. The power and validity of this study is then shown by providing useful examples of where these disciplines have been applied together previously. These previous studies provide a baseline for the application of logistic regression to conflict prediction as well as testing model accuracy.

III. Methodology

3.1 Overview

This chapter details the methods used to apply conflict prediction to the Arab Spring. This section starts with an understanding of the scope of this study and the assumptions necessary to answer the research questions. The next section describes the database used to explore significant factors in this study and a description of the dependent variable including how it differs from previous studies to better relate to the Arab Spring's specific application of conflict prediction. Because the database includes variables from multiple open source archives and due to the fact that many of these archives' updates lag by two years, there were data gaps in many of the database's variables. To overcome this challenge, the next section outlines a method for imputing variables from the database for use in future studies addressing conflict prediction. Following the assembly of completed variables, the next section outlines relevant independent variables to be referenced through the rest of the study. Next is a discussion of the creation of principal components and trend variables to refine the model building processes, and then address the steps and outcomes of the four model building processes, PSOC, LSOC, PCR, and RPCR.

3.2 Assumptions

As noted by Boekestein [3] and Shallcross [4], this study assumes that the candidate, independent variables represent true country information, that they can show the tendency of a nation to tend toward or away from conflict, and that those independent variables found to be significant to predicting conflict levels remain significant over the entire time span of analysis. This study also assumes that the region definition identifies all countries affected and no countries unaffected by the Arab Spring.

Another assumption to the study of the Arab Spring is that an initial conflict shock from increased protests in a region creates a unique conflict scenario that cannot be accurately predicted using normal conflict models.

3.3 Limitations

The greatest limitation this study faced is caused by the reality of the environment it tries to represent. The use of 60 observations (12 countries x 5 years) limited statistical processes when analyzing a nearly equivalent number of variables. Some aspects of the analysis had to rely on correlation analysis and indications from research instead of standard logistic regression techniques to continue model building processes at certain steps. Also, some aspects of the research hypothesis could not be directly tested due to variable inadequacies for some countries as outlined in the focus on imputation.

3.4 Database Description

The development of the set of candidate, independent variables began with AFIT's PMESII Internal Conflict Database which combines variables describing different attributes under the PMESII (Political, Military, Economic, Social, and Infrastructure and Information Systems) structure as defined by the United States Department of Defense's Joint Publication 2-01.3, Joint Intelligence Preparation of the Operational Environment [29]. This database includes data from over 182 countries and multiple data sources. This study sifted through the data to select variables with at least some values present for the Arab Spring countries and reduced redundant variables. This resulted in 225 candidate variables along with the HIIK conflict indicator to be altered as the dependent variable. Seventy-nine of the 225 variables were deemed useful for this study based upon research and further redundancy limiting. This step was necessary to limit extreme multicollinearity issues and increase interpretability of the eventual models. Following this reduction, the candidate variables were prepared for imputation.

3.5 Imputation

Missing Data

The initial data set compiled from the PMESII database had missing values for the selected country-year observations in 48 of 79 variables. Twenty-six of the incomplete variables were missing all observations for at least one of the countries. Under the assumption that the magnitude of each country's observations should be independent of other countries, foregoing regional time trends, imputation was not possible for variables missing an entire country's observations. Because of this, these variables were not considered in model building and were removed from the data set. This left a data set with 4.51% missingness. Nine of the remaining incomplete variables were missing all occurrences from 2015, the most recent year analyzed. It was determined that these values could be imputed however from the previous years' data as well as from using predictive methods from other complete variables for 2015 that we will discuss later in this chapter.

The variables missing 2015 data as well as those missing data in a less observable pattern follow the classification of data missing at random (MAR). This distinction was made as the missingness of the data could be attributed to other variables in the dataset, and not due to the value of the missing data themselves as is commonly seen in survey data [30]. In this study's data, missingness seemed to be caused by either the year variable, or political turmoil, accounted for by other variables, that made it difficult to gather that data. An example of this is seen in the observations of Consumer Price Index (CPI) and Inflation for Syria. There are no observations for either of these variables from 2013 to 2015. This is most likely due to the fact that Syria's political structure, and therefore ability to record aggregate, economic data, was impacted by the civil war in Syria and occupation of the Islamic State [17].

Multiple Imputation by Chained Equations

When considering the MAR observation deletion was not considered due to the already small sample size of five data points per country. Instead, multiple imputation by chained equations (MICE) was used according to the MICE package in R [31]. Multiple imputation is a method that can apply several types of relevant imputation methods, such as linear regression or classification trees, to each variable with missing data. Using the specified methods, it calculates m completed data sets through Monte Carlo Simulation to capture the variability and uncertainty of the missing data where m is set by the analyst and is classically set to ten [32]. Most applications of MICE suggest keeping all m sets throughout the analysis and pooling the results at the end of the study as this has been shown to maintain the most realistic variability, however, in this case, m imputed data sets were pooled prior to analysis as the low rate of imputed data (4.51%) should not produce a meaningful difference throughout the analysis, and a post-analysis pooling is not conducive to the model building testing methodology used here [32]. Although it is held by Schafer [33] that the premature pooling of imputations may cause a slight bias, using the mean of the distribution of estimates still allows for some of the variability of the missing data as well as the possibility of non-normal distributed observations. These attributes are often lost in single imputation techniques which provide inflated precision estimations. The methods of imputation compared for use in the remaining variables are shown in Table 4.

Table 4. Imputation Methods Used

cart	Classification and regression trees
pmm	Predictive mean matching
norm	Bayesian linear regression
rf	Random Forest
mean	unconditional mean
locf	Last observation carried forward
research	Polity missing values were all transitional, scaled to 0; If no troops deployed, no value given for Deployed US Troops

Eighteen of the 23 remaining variables with missing data required the use of the MICE package in R. Although one of the benefits of the MICE package is that it contains methods for imputing both continuous and discrete variable types, all of the variables requiring imputation in R were continuous so only applicable methods were analyzed. The first three methods in Table 4 denote the methods that were used in the final imputations in R. The cart method runs a classification or regression tree depending on the variable under consideration and may have been favored in the testing process due to its ability to incorporate interaction terms in the prediction of the missing data. For further discussion of MICE's cart details, reference Doove et al. [34]. The second method, pmm, "imputes missing values by means of the nearest-neighbor donor with distance based on the expected values of the missing variables conditional on the observed covariates" [35]. Vink [35] tests MICE's pmm method to other similar imputation techniques and favors pmm in a multivariate simulation experiment due to its low bias to the tested data. Bayesian linear regression is performed for the norm method and uses a posterior probability distribution from linear regression that is conditioned on the observed data [31]. These three methods, as well as random forest and the unconditional mean, were compared for each of the 18 variables using the mean of $m = 10$ simulated data sets. Imputations were created using data from the non-missing values of the given variable and the complete variables in our data. Through experimentation, a single method for each variable was used, however random forest and unconditional mean were not found to be the preferred strategy for any of the incomplete variables. The preferred strategy was based on the method that produced imputations most closely distributed to the observed data within each country as described in the following subsection.

Testing Imputation Methods

The five methods in R were compared statistically through the use of the non-parametric, 2-sample Anderson-Darling (A-D) and Kolmogorov-Smirnov (K-S) tests. Abayomi et al. [36] cover several methods for analyzing multiple imputations and cite the use of the K-S test as a technique to determine if imputed values are similar to observed data given the assumption that the imputed data should follow the same distribution. The comparison of the K-S and A-D tests was performed after reading the findings of Engmann and Cousineau [37] which support the A-D test as a dominant test to K-S for observing small differences in distributions at the tails as well as other moments of the distribution. Because of the posit that the A-D test provides a stricter test for determining the likeness of two samples, the findings of the A-D test were used as the main distinction between methods. The K-S test was also used in this analysis to provide a more conservative comparison between imputed and original data since the imputed data may not fit exactly into the distribution of the original data especially for variables missing an entire year's worth of data. Both of these tests use the null hypothesis that the two samples do, in fact, come from the same parent distribution. A low p-value of the tests indicates that the imputed data and the non-missing data do not come from the same parent distribution.

Transformation of Data Prior to Imputation

As mentioned earlier, the data requires the assumption that the magnitude of each country's observations are independent of each other. Because of this, imputations had to be compared within countries. This proved to be difficult as some countries were missing only one value for the given variable making a non-parametric, 2-sample test impossible. Instead, a statistic was developed according to Equation 10 that normalized the observed data prior to imputation.

$$
\begin{gathered}
Z_i = \frac{X_i - \overline{X}_c}{\sigma_c} \\
i = observation \in 1, 2, ..., n \\
c = country \in 1, 2, ..., 12
\end{gathered}
\tag{10}
$$

The normalization ensured that the preferred method created imputations from the same parent distribution as the corresponding observed values within each country. Key to this approach is the assumption that each variable is distributed according to the same family of distributions while with possibly different distribution parameters. Imputation could be run for all countries within a variable simultaneously under the assumption that all countries come from the same family of distributions as they should all trend similarly during observed years. This allowed the use of non-parametric testing for all variables and increased the statistical insights by increasing the number of observations used in imputation from $5 - (\# \; of \; missing \; values \; for \; country \; c)$ to $60 - (\# \; of \; total \; missing \; values)$.

The efficacy of the test on the transformed variables can be understood through the comparison of two hypothetical distributions from Country A and Country B. Country A is sampled with observations for variable i that follow a skewed Weibull distribution with mean, $\overline{\mu}_1$ and standard deviation, $\overline{\sigma}_1$. Country B is sampled with observations for variable i that follow a skewed Weibull distribution with mean, $\overline{\mu}_2$

and standard deviation, $\overline{\sigma}_2$. Once normalized, both samples would follow the same distribution with values $\frac{X_a^i - \overline{X}_a}{\overline{\sigma}_a}$ and $\frac{X_b^i - \overline{X}_b}{\overline{\sigma}_b}$ for Countries A and B, respectively. From there, imputation would be performed on the normalized samples from all countries. The imputation produces a normalized value with realistic variation from the common family mean and standard deviation based on the observed data for that country-year. The values are then transformed to the country of interest distribution. Solving for the X_1^i and X_2^i would produce observations scaled to the distributions of Countries A and B separately according to $\overline{\mu}_1$ and $\overline{\sigma}_1$ for Country A and $\overline{\mu}_2$ and $\overline{\sigma}_2$ for Country B as shown in Equation 11.

$$[H]\, X_i = Z_i * \sigma_c + \overline{X}_c \tag{11}$$

Without normalizing the data within each country, the imputed data could artificially be considered to fit the given data distribution well due to the spread in magnitude of countries' values within a variable. The distribution of data without normalization is multimodal with peaks near each country's mean. This would cause imputations from the preferred strategy to fall near each of the country peaks and not the country distribution that the imputed observation belongs to instead.

A-D and K-S Test Results

The results from the A-D and K-S tests using the normalized statistics for all five methods are shown in Tables 5 through 8.

Table 5. AD and KS Test p-values for each Strategy

		Arable Land	Armed Forces	Birth Rate	Computer Exports
A-D	cart	0.4436	0.0132	0.0449	0.0950
	pmm	0.8676	0.4061	0.2487	0.0033
	norm	0.1724	0.1177	0.0154	0.2193
	rf	0.0762	0.0352	0.0568	0.0608
	mean	0.0015	0.0005	0.0011	0.0040
K-S	cart	0.6476	0.0129	0.0348	0.1100
	pmm	0.8186	0.4749	0.3291	0.0014
	norm	0.6476	0.4499	0.2165	0.8186
	rf	0.2165	0.0303	0.0468	0.0601
	mean	0.0068	0.0048	0.0033	0.0043

Table 6. AD and KS Test p-values for each Strategy

		CPI	Death Rate	Fertility Rate	HDI	Imports	Inflation
A-D	cart	0.2642	0.8899	0.1359	0.6476	0.2389	0.5291
	pmm	0.6314	0.1725	0.7358	0.0468	0.0834	0.0167
	norm	0.4323	0.1359	0.3291	0.0348	0.1694	0.3447
	rf	0.1072	0.2685	0.0348	0.0621	0.1694	0.3447
	mean	0.0473	0.0068	0.0047	0.0068	0.0834	0.0167
K-S	cart	0.2755	0.7419	0.0497	0.6752	0.2519	0.5578
	pmm	0.6336	0.1552	0.4626	0.0659	0.1820	0.0571
	norm	0.6934	0.0510	0.0320	0.0105	0.2090	0.3701
	rf	0.0758	0.1201	0.0520	0.0890	0.2090	0.3701
	mean	0.1079	0.0023	0.0013	0.0022	0.1820	0.0571

Table 7. AD and KS Test p-values for each Strategy

		Tourism Expenditure	Internet Users	Female Youth Labor	Youth Labor
A-D	cart	0.1297	0.7635	0.5449	0.1537
	pmm	0.0000	0.8310	0.0009	0.0013
	norm	0.0135	0.8356	0.1978	0.1108
	rf	0.0135	0.8356	0.1978	0.1108
	mean	0.0000	0.8310	0.0009	0.0013
K-S	cart	0.2928	0.8232	0.5596	0.2165
	pmm	0.0015	0.9402	0.0022	0.0047
	norm	0.0180	0.9617	0.2685	0.0621
	rf	0.0180	0.9617	0.2685	0.0621
	mean	0.0015	0.9402	0.0022	0.0047

Table 8. AD and KS Test p-values for each Strategy

		Life Expectancy	Merchandise	Refugee Asylum	USAID Econ
A-D	cart	0.2387	0.4155	0.0837	0.6854
	pmm	0.0015	0.8266	0.4699	0.8126
	norm	0.0302	0.8266	0.7043	0.8126
	rf	0.0302	0.8266	0.7043	0.8126
	mean	0.0015	0.8266	0.4699	0.8126
K-S	cart	0.2685	0.5260	0.1385	0.7355
	pmm	0.0068	0.9273	0.5821	0.8974
	norm	0.0621	0.9273	0.7805	0.8974
	rf	0.0621	0.9273	0.7805	0.8974
	mean	0.0068	0.9273	0.5821	0.8974

The maximum test value for all variables passed above the 0.1 alpha level. This shows that MICE is an acceptable technique for creating imputations that closely follow the given data. Eight of these maximum test p-values were calculated using pmm, eight using cart, and two using norm. For all variables except CPI, the A-D and K-S tests agreed on the best method for imputation. The CPI K-S test p-values for norm and pmm were similar at the 0.05 level so the determination was deferred to the A-D test outcome, pmm. Ties in p-values for Merchandise Imports and USAID Econ defaulted to pmm due to the research from Vink et al. [35]. After testing each imputation, all values were transformed back to their in-country, magnitude-based values through Equation 11.

Practical Check of Imputations

All imputations were visually checked in the full data set to ensure that extreme changes or lack there-of made sense in each country-year observation. One of these logic checks was made after viewing the following results of Syria's inflation scores. The data are seen in Table 9.

Table 9. Syrian Inflation Imputation Comparison

Country	Year	Inflation
Syria	2010	4.397414
Syria	2011	4.753164
Syria	2012	36.7023
Syria	2013	21.97421
Syria	2014	20.03748
Syria	2015	20.35689

The highlighted numbers denote the imputed values. It may appear at first that the outlier of 2012 as compared to earlier years causes the imputations for 2013-2015 to be incorrectly high, however, Trading Economics indicates that for one month in 2013, Syria experienced an inflation rate of 121.29% [38]. This provides evidence to retain these values as valid imputations for inflation. There were no other alarming deviations from observed data or researched history. The data then passed both the statistical and practical tests and was ready for further analysis. This supports the MICE approach as a possible answer to research question S1: what is the best method for imputing missing PMESII data?

Imputation of Variables Outside of R

Five variables required imputation methods other than multiple imputation in R. Contract Time, Business Time, and Tax Time were non-dynamic factors so the most recent observation for each country was carried over to the missing values. Polity was only missing observations where countries were experiencing anarchy, foreign interruption, or transition periods so these values were scaled to zero as outlined in Section 3.7. US Troops Deployed was only missing data for observations where there were no troops deployed to that nation during that year so these observations were imputed with zero as well. The method chosen for each variable imputed is shown in Table 10.

Table 10. Imputation Methods Used for Each Variable

Variable	Method
Arable Land	pmm
Pct Armed Forces	pmm
Birth Rate	pmm
CPU Exports	norm
CPI	pmm
Death Rate	cart
Fertility Rate	pmm
HDI	cart
Imports	cart
Inflation	cart
Tourism	cart
Internet Users	norm
Female Labor Participation	cart
Youth Labor Participation	cart
Life Expectancy	cart
Merchandise Imports	pmm
Refugee asylum	norm
USAID	pmm
Contract Time	locf
Business Time	locf
Tax Time	locf

Deployed US Troops	research
Polity	research

3.6 Dependent Variable

As mentioned in Section 2.4 the dependent variable is a variation of previous conflict studies' HIIK conflict indicators. To fully understand the meaning of this difference, a description of how the HIIK computes their six-level conflict barometer is necessary. The HIIK defines political conflict as, "a positional difference, regarding values relevant to a society - the conflict items - between at least two decisive and directly involved actors, which is being carried out using observable and interrelated conflict measures that lie outside established regulatory procedures and threaten core state functions, the international order or hold out the prospect to do so" [1]. They explain that along with conflict items, actors, and measures, political conflicts have the property of intensity as defined by Figure 3.

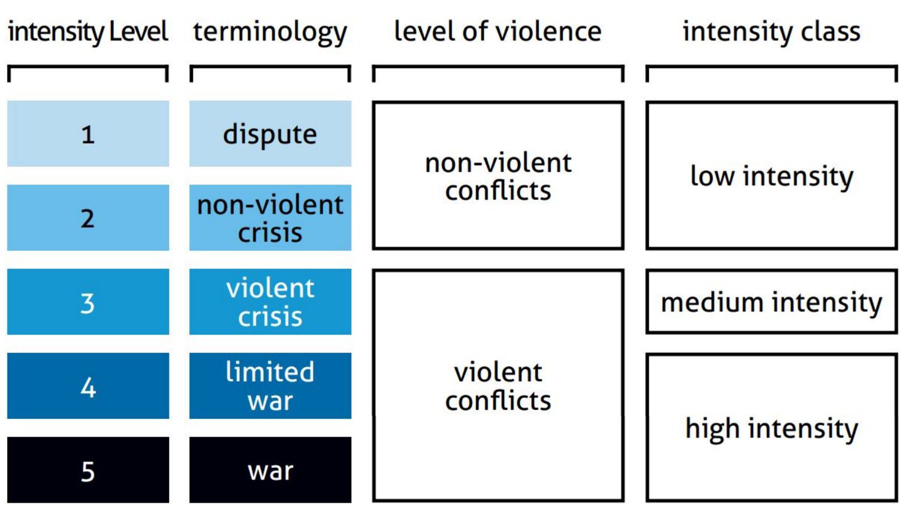

Figure 3. The Concept of Conflict Intensity [1]

Countries are classified into one of the five intensity levels illustrated in Figure 3 or no conflict based on five classes of identifiers: weapons, personnel, casualties, refugees and IDPS, and destruction. Each identifier is rated on a scale of 0, 1, or 2 with an increase in the scale equal to an increase in conflict intensity rating. Each indicator receives a score according to Figures 4 through 8.

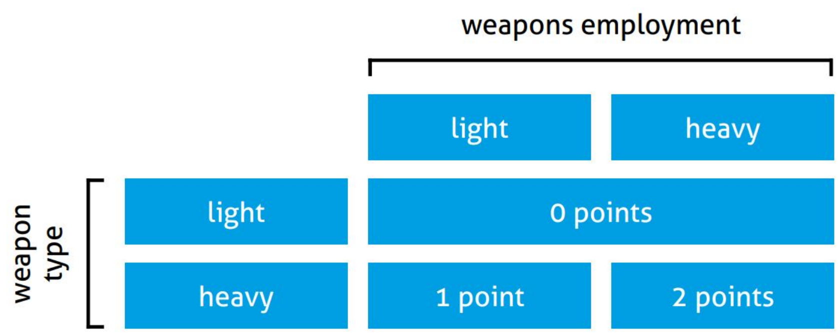

Figure 4. Weapons Criteria

low	medium	high
≤ 50	> 50 and ≤ 400	> 400
0 points	1 point	2 points

Figure 5. Personnel Criteria

low	medium	high
≤ 20	> 20 and ≤ 60	> 60
0 points	1 point	2 points

Figure 6. Casualty Criteria

low	medium	high
≤ 1 000	> 1 000 ≤ 20 000	> 20 000
0 points	1 point	2 points

Figure 7. Refugee and IDP Criteria

low	medium	high
within 0 dimensions	within 1 - 2 dimensions	within 3 - 4 dimensions
0 points	1 point	2 points

Figure 8. Destruction Criteria

The scores are then conglomerated using the structure seen in Figure 9.

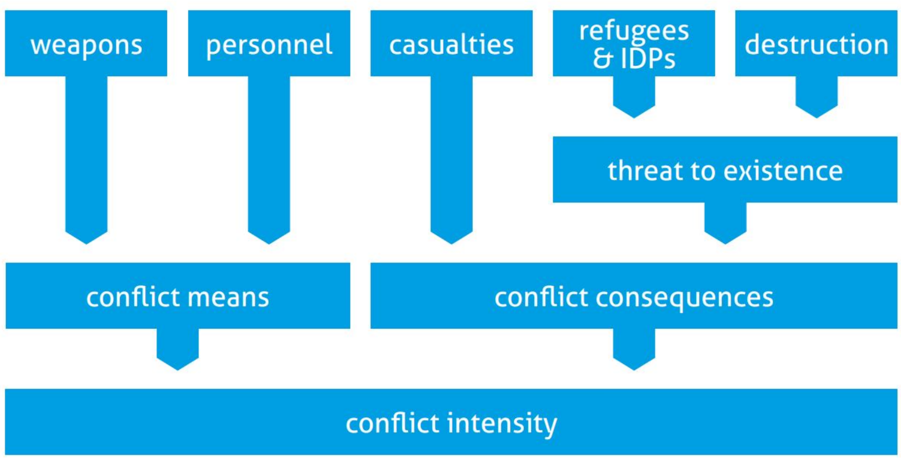

Figure 9. Indicator Criteria Tree

The final intensity rating is given using the table in Figure 10.

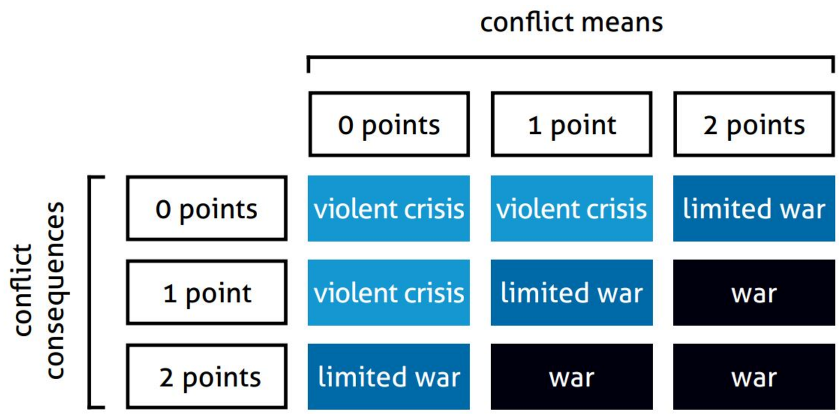

Figure 10. HIIK Rating Table

The level of protest that defined the beginning of the Arab Spring forced all affected nations to a HIIK indicator level of 3 or higher indicating a violent crisis, limited war, or war. Table 11 supports this hypothesis with all countries in our study achieving a HIIK conflict intensity rating of three or higher.

Table 11. 2011 HIIK Conflict Intensity Levels

Country	Year	HIIK Highest Level of Conflict
Bahrain	2011	3
Egypt	2011	5
Jordan	2011	3
Kuwait	2011	3
Libya	2011	5
Mauritania	2011	4
Morocco	2011	3
Oman	2011	3
Saudi Arabia	2011	3
Syria	2011	5
Tunisia	2011	4
Yemen	2011	5

To achieve the desired study outcomes, this analysis differentiates between countries that will experience "high-intensity, violent conflict" and those that will not [1]. According to Figure 3, this would include HIIK levels four and five. Previous studies included HIIK level three in their in-conflict definition however we hypothesize that nations in a "violent crisis" would not garner US intervention which was the case for several of the countries in this study [3, 4, 5]. Although a violent crisis year may eventually lead to higher levels of conflict we are only interested in predicting which countries will attain the higher levels giving US government decision makers an improved ability to intervene in environments with large threats to human life. Recent studies in Shallcross [4] and Leiby [5] have also used conditional transition dependent variables instead of a pure conflict indicator. Their studies were concerned with predicting transitions into conflict as well as transitions out of conflict. This study takes a more conservative approach to flag all occurrences of high-intensity, violent conflict. The dependent variable in our analysis is defined by the following probability calculation.

$$P(Y_i > 3|Y_{i-1} \leq 3) \cup P(Y_i > 3|Y_{i-1} > 3) = P(Y_i > 3) \tag{12}$$

This captures both country-year observations that went from a lower level of conflict to high-intensity, violence and those that started at this higher stage of conflict and remained so from the previous year. A conditional transition dependent variable would be unable to capture countries such as Yemen that had a HIIK rating over three both before and after the Arab Spring as being a negative (1) event. It also would not be able to incorporate the multi-stage behavior of the conflicts seen in the Arab Spring consisting of violence, government push-back and then multi-factional fighting over control of the nation which might take multiple years to occur.

3.7 Relevant Independent Variable Descriptions

This section provides a brief explanation of the variables used in the four model building techniques with a description of their origin, definition, and any transformations applied. Reference Appendix A for a list of all independent variables considered in modeling.

Consumer Price Index

Consumer Price Index (CPI) is a cost index with base year 2010. The data originates from the International Monetary Fund, and was collected via The World Bank [39]. According to the IMF, CPI measures “the rate at which the prices of consumer goods and services are changing over time” [40].

Human Development Index

The Human Development Index (HDI) as a factor created by the United Nations Development Program (UNDP) [41]. HDI tries to capture a country’s level of development based upon citizens’ lives, not macroeconomic indicators as commonly used. To do this, it aggregates subcomponents of three main indicators of country development: Long and Healthy Life, Knowledge, and A Decent Standard of Living. The first indicator is represented by the variable Life Expectancy at Birth. The second is the arithmetic mean of the mean number of years of schooling completed for the population 25 and older and the expected number of years of schooling for children about to begin schooling. The third is represented by the logarithm of Gross National Income per Capita. All four of the subcomponents are normalized at their base level and then are aggregated using the geometric mean fitting the variable on a [0,1] scale.

Voice and Accountability: World Governance Indicator

The Voice and Accountability Percentile Rank is one of six World Governance Indicators created by the World Bank that attempts to capture the population’s perception of their ability to affect their government or freedom [42]. It is calculated by first scaling multiple factors from outside sources to a [0,1] scale pertaining to freedom and citizen effect on government and then aggregating these factors as a linear combination with weights determined by an Unobserved Components Model. Countries are then scaled as a percent between the minimum and maximum scores worldwide. The Voice and Accountability Rankings for 2011 are depicted in Figure 11.

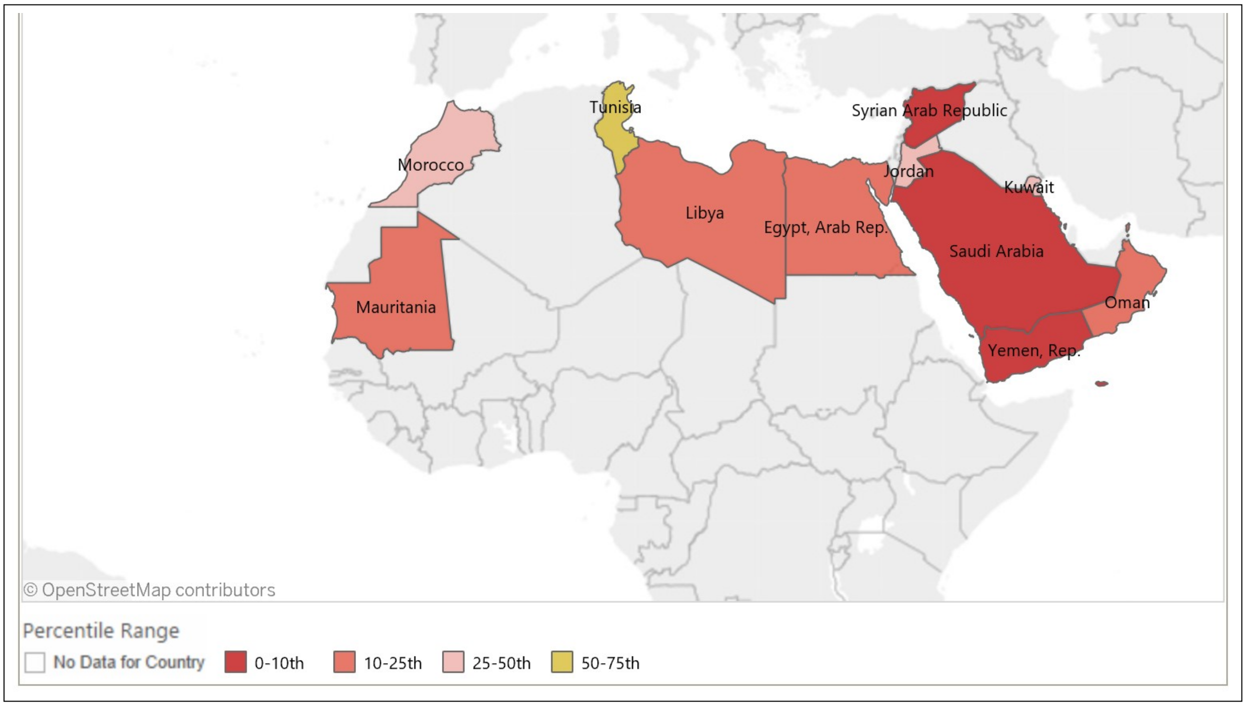

Figure 11. Voice and Accountability Rankings in Arab Spring Countries [42]

International tourism, expenditures (current US$)

International tourism expenditures (Tourism) is a measure collected by the UN World Tourism Organization and was accessed via The World Bank [39]. This measure provides the amount of money in US dollars that the nation's population has spent on tourism in other countries.

Polity

Revised Combined Polity Score for Time-Series Analysis is a variable from the Center for Systemic Peace to rate countries' government regimes on a scale from "hereditary monarchy" at -10 to "consolidated democracy" at 10 [43]. The Center of Systemic Peace suggests this spectrum of regimes incorporates "incoherent, authority regimes" as a zero, but classifies regimes experiencing "foreign interruption", "anarchy", and "transition" as separate from the [-10, 10] scale and instead at values of -66, -77, and -88 respectively. This study transforms this variable, Polity, to incorporate all observations in these extreme cases as zero as these three cases fall under similar definitions as those normally with a zero rating. This transformation is more significant in the univariate case with this study's dependent variable than any previous derivation of the Polity variable [4]. The base score for non-extreme cases is calculated by allocating points to a country's democratic and autocratic rating based on five categories then subtracting the autocratic rating from the democratic rating. The five categories are Competitiveness of Executive Recruitment, Openness of Executive Recruitment, Constraints on Chief Executive, Regulation of participation, and Competitiveness of Participation.

Corruption Perception Index

The Corruption Perception Index (Corruption) by Transparency International attempts to enumerate corruption in the public sector as described by "analysts, businesspeople and experts in countries around the world" [44]. As such, CPI relies on both survey results as well as other outside-sourced variables. It includes the following 13 variables in its calculation of Arab States:

Table 12. Variables Incorporated in Corruption Perception Index

Variable Name
World Bank CPIA
World Economic Forum EOS
Global Insight Country Risk Ratings
Bertelsmann Foundation Transformation Index
African Development Bank CPIA
IMD World Competitiveness Yearbook
Bertelsmann Foundation Sustainable Governance Index
World Justice Project Rule of Law Index
PRS International Country Risk Guide
Varities of Democracy Project
Economist Intelligence Unit Country Ratings
Freedom House Nations in Transit Ratings
PERC Asia Risk Guide

All variables are standardized using the z-score method and are then scaled from 0-100. The means of all variables are then averaged to compute the final index value.

Percent Border Conflict

Percent Border Conflict (Pct BC) was originally calculated by [3]. It attempts to represent the effect that neighboring countries' conflict has on each other by creating a linear combination of countries' HIIK conflict level and the percent of its border shared with the corresponding country as seen in Equation 13.

$$Pct\ BC = \sum_{i=1}^{b} x_i p_i \tag{13}$$

$$b = \#\ of\ bordering\ nations$$

$$x_i = HIIK\ intensity\ level\ for\ nation\ i$$

$$p_i = \%\ of\ border\ shared\ with\ nation\ i$$

Mobile cellular subscriptions (per 100 people)

This variable, named Mobile Cell, originates from the UN's International Telecommunication Union and was accessed via The World Bank [39]. It takes the total number of mobile cellular subscriptions in the country and relates it to the total population to determine the average number of subscriptions per 100 people.

Armed forces personnel (% of total labor force)

This variable, referred to as Pct Armed Forces, was accessed via The World Bank [39]. According to Trading Economics, this factor considers personnel in the armed forces to include "active duty military personnel, including paramilitary forces if the training, organization, equipment, and control suggest they may be used to support or replace regular military forces" [45].

Labor force participation rate for ages 15-24, total (%) (modeled ILO estimate)

This variable is referred to as Youth Labor Participation. It originates from the International Labor Organization and was accessed from The World Bank [39]. Differing from unemployment it looks at the "proportion of a country's working-age population that engages actively in the labour market, either by working or looking for work" [46].

Population ages 0-14 (% of total)

To standardize with Shallcross [4], this variable is called Youth Bulge according to the research from Urdal [19]. This variable originates from The World Bank and is estimated with data from United Nations Population Division's World Population Prospects [39].

Refugee population by country or territory of origin

This variable is referred to as Refugee Origin. It originates from the United Nations High Commissioner for Refugees (UNHCR) and was accessed from The World Bank. This variable is a count of the number of people that are considered in refugee status from a given country in the given year. The UNHCR considers refugees to be persons "fleeing conflict or persecution" [47].

USAID Economic Assistance ($ US)

USAID Economic Assistance, which is shortened to USAID, is the total amount of financial assistance given to countries by the U.S. Agency for International Development for the development of countries to promote diplomacy and further US interests in the international community [48].

Deployed US Troops

Deployed US Troops shows the number of troops deployed to each country for each year [39]. This variable was gathered from The World Bank.

Uneven Development

The Uneven Development variable is one of twelve Fragile States Indices by the Fund for Peace [49]. It measures the economic disparages between different groups including racial, ethnic, and religious as well as the ability of different groups to improve their economic status. All of the Fragile State Indices are calculated by aggregating inputs relative to the variable's theme from media database searches, open-source data, and social scientist inputs.

Government Effectiveness: Percentile Rank

The Govt Effectiveness variable is another of the World Governance Indicators that captures the population's perception of the "quality of public services," degree of civil service's "independence from political pressures," "quality of policy formulation and implementation," and "credibility of the government's commitment to such policies" [42].

Fertility rate, total (births per woman)

Fertility rate was obtained through the World Bank which aggregated the statistic from multiple sources [39]. It shows the average number of births for women of all ages in a country.

Three-Year Trend Variables

Consistent with previous conflict prediction studies, trend variables are included in the model building strategies to determine if previous states of variables have caused a degradation to stability and increased conflict probability. The three-year trend variables were included for all variables in the base dataset using Equation 14. All trend variables are referred to as 3YT and the variable name.

$$3 - Year\ Trend = Variable\ Level_y - Variable\ Level_{y-3} \tag{14}$$

3.8 Principal Component Analysis

The method introduced in Section 2.3 was used to obtain the PCs in the models. Both the component loadings as well as the component scores were calculated for further analysis. The component loadings are the correlation coefficients between the principal components and original variables. These allow interpretation of the meaning of the principal components. The loadings were calculated using all 84 variables and trends according to Equation 15.

$$Loading\ Matrix = \left(\sqrt{diag(\rho(Z))}\right)^{-1} V\sqrt{D} \tag{15}$$

Where $\rho(Z)$ is the correlation matrix of the standardized data, V is the matrix of eigenvectors corresponding to each component, and D is a diagonal matrix with diagonal values corresponding to eigenvalue of each component, λ_j. The eigenvalue for each component is calculated according to Equation 16 where I is the identity matrix corresponding to the number of variables in Z and v_j is the eigenvector corresponding to eigenvalue λ_j.

$$(Z - \lambda_j I)v_j = 0 \tag{16}$$

Next, the PCA scores were calculated which map the principal components to all observations. These were calculated via Equation 17.

$$Component\ Scores = Z\left(\sqrt{diag(\rho(Z))}\right)^{-1} V \tag{17}$$

V in this case is the matrix of eigenvectors for all principal components. As stated earlier, the number of components considered in the models was not reduced due to the fact that PCR may find components that explain a small amount of the variation of Z, but a large amount of the variation in the dependent variable. PCA resulted in 84 components to be considered in the modeling stage. These components should improve analysis of the multiple collinear PMESII variables.

3.9 Model Building Strategies

This section provides an overview of the outcomes of the four model building methods. The test of these methods is a significant contribution to the conflict prediction line of study as multiple modeling techniques are compared which investigated a large number of independent variables. As the general steps for each technique were previously explained in Section 2.3, only the steps pertinent to validate the statistical relevance of the proposed processes and resulting models are included.

Purposeful Selection of Covariates

The PSOC model was initialized by testing the univariate significance of all 84 base variables and their corresponding 3-year trend variables. Thirty-four of these variables passed the Pearson χ^2 test at the 0.25 α level. An initial, multivariate model was produced from these 34 variables. This model was unstable due to high multicollinearity amongst many of the included PMESII variables. Before continuing the steps of purposeful selection, a new, full model, with decreased multicollinearity, was needed to determine how to reduce the full model. This endeavor began by inspecting the correlation between all 84 variables as depicted in Figure 12. This figure shows the correlation of all 84 variables with each other. Large, blue, opaque circles indicate variables with high, positive correlation and large, red, opaque circles indicate variables with high, negative correlation. Intersections with small, translucent circles indicated low levels of correlation that did not signal possible multicollinearity.

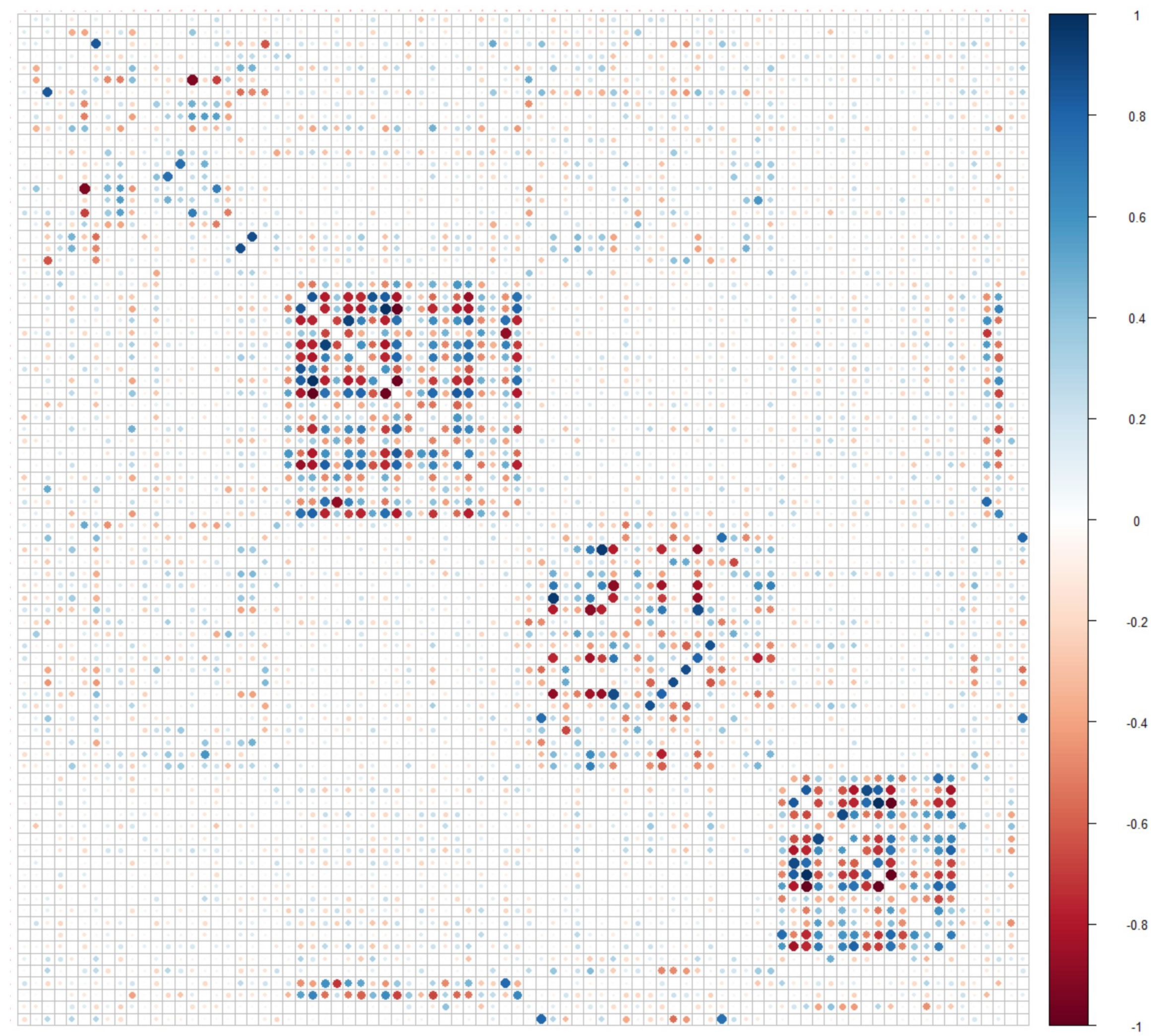

Figure 12. Correlation of All Base and Trend Variables

Since it is preferable to make the PSOC model as statistically based as possible, our methodology eliminated correlated variables based upon the lack of strength of their univariate significance with the dependent variable. For groups of variables with correlation greater than 0.5, the variable with the highest univariate Pearson χ^2 test p-value was removed from the model. This elimination strategy was repeated until the model was stable, reducing the size of the full model from 34 variables to 17 variables. From this point there was still a large amount of collinearity among the independent variables apparent as none of the parameters had significant Wald statistics in the reduced model proving further reduction was necessary. Variables were again removed from the model to form a reduced model, now based on the significance of the Wald test in the multivariate model. The variable with the highest Wald statistic p-value was removed one at a time until all variables' p-values were below 0.05. This resulted in the three-variable reduced model in Figure 13.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	23.797105	3	47.59421	<.0001*
Full	16.583595			
Reduced	40.380700			

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	56	16.583595	33.16719
Saturated	59	0.000000	Prob>ChiSq
Fitted	3	16.583595	0.9935

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	1.91029574	0.7004873	7.44	0.0064*
3YT CPI	-0.2274549	0.0674159	11.38	0.0007*
3YT HDI	233.682334	71.944195	10.55	0.0012*
Intnl Tourism per billion	0.31201803	0.1532337	4.15	0.0417*

For log odds of 0/1

Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
3YT CPI	1	1	30.2591453	<.0001*
3YT HDI	1	1	31.2337263	<.0001*
Intnl Tourism per billion	1	1	5.18445212	0.0228*

Figure 13. PSOC Reduced Model

Due to the amount of collinearity in the full model however, there were large changes in the parameter estimates (>20%) between the full and reduced models which suggests that some factors may be confounding and are needed to adjust the effect of certain parameter estimates. The next step prescribed added all previously removed variables from the full model back into the reduced model one at a time to test this. The change in parameter estimates could not be reduced below 20% without forcing all variables in the reduced model to be insignificant. A loss of overall accuracy due to missing confounding factors is accepted as it outweighs the loss of interpretability of a model with a large amount of non-significant variables. The conclusion is to retain the original, reduced model of three variables. The loss of accuracy is seen in the outcome of the Likelihood Ratio Test between the full and reduced models which results in a χ^2 test with p-value of 9.5E-5. This low p-value indicates that the reduced model is most likely not as accurate as the full model.

Considering the reduced model, the next step looked to add variables into the model that were insignificant in the univariate tests. This is another attempt to incorporate confounding factors that alter the parameter estimates to more accurately model the behavior of the dependent variable. After testing all variables that were not included in the initial, full model, two variables, 3YT Refugee Origin and Voice and Accountability, were significant at the 0.05 alpha level when added to the reduced model. In both instances of adding these variables into the reduced model however, International Tourism's parameter estimate became insignificant. An investigation followed, replacing International Tourism with the newly significant variables included and found that the model with 3YT CPI, 3YT HDI, and 3YT Refugee Origin had the most significant variables and the largest χ^2 statistic of overall model fit. This model is accepted for the next step of PSOC. All three variables did not violate the logit linearity assumption using plots of the independent variables versus the logit transformation of the predicted values as seen in Figure 14.

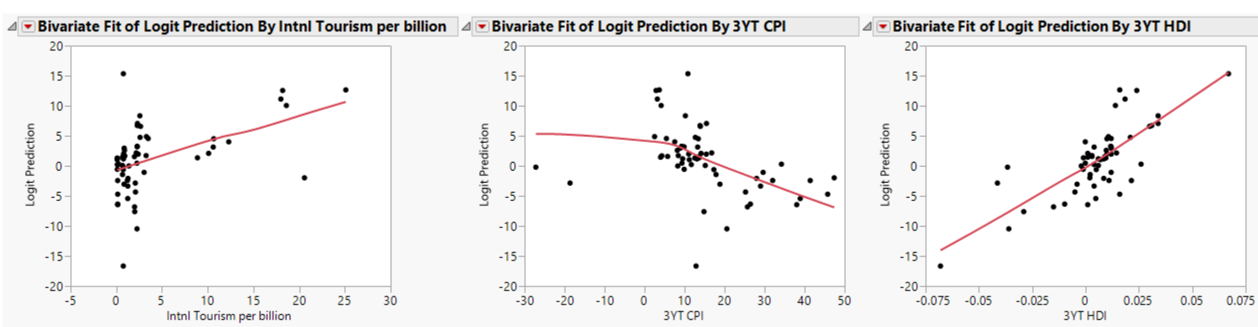

Figure 14. Plot of Reduced Model Variables vs. Logit Transformation of Predicted Values

The final step tested for significant interactions between the variables included in the model, but did not find any that significantly improved the model. The final model leading into validation in Chapter IV is seen in Figure 15.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	27.207368	3	54.41474	<.0001*
Full	13.173332			
Reduced	40.380700			

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	56	13.173332	26.34666
Saturated	59	0.000000	Prob>ChiSq
Fitted	3	13.173332	0.9997

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	3.61189251	1.1418422	10.01	0.0016*
3YT CPI	-0.2952263	0.086702	11.59	0.0007*
3YT HDI	334.924378	104.50275	10.27	0.0014*
3YT Refugee Origin/1000	-0.001315	0.000491	7.17	0.0074*

For log odds of 0/1

Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
3YT CPI	1	1	36.2611068	<.0001*
3YT HDI	1	1	40.2760887	<.0001*
3YT Refugee Origin/1000	1	1	12.0049766	0.0005*

Figure 15. Final Model For PSOC

PSOC represents a statistical and uninformed perspective on conflict prediction. It will help retain any variables missed by background research and may lead to a more general approach for repeating this study for other regions if a new conflict shock occurred in another region or under different circumstances. The flow of variables into and out of the model is summarized in Table 13 to demonstrate the benefit of analyzing the inclusion of variables at multiple stages of model building.

Table 13. Flow of Variables into and out of Model in PSOC

Step	Number of Variables
1	84
2a	34
2b	17
2c	3
3a	14
3b	3
4a	4
4b	3
5	3
6	3

Logical Selection of Covariates

The LSOC model tried to repeat most of the same methodology as that of PSOC but limited the original pool of variables to be analyzed based upon our hypothesis from Section 2.2. To do this, the candidate, independent variable pool was limited to relevant variables to the nine categories seen in Table 14.

Table 14. Categories Considered for LSOC Variable Inclusion

Autocracy
Corruption
Protest Threat
Opportunity
Quality of Life
Violent Response
Regional Effect
Technology
Youth Bulge

The first six of these are explained in the anecdotal hypothesis in Section 2.2, and the last three are based on the hypotheses tested by Costello et al. [6] and Shallcross [4]. Fifteen base variables as well as their corresponding trend variables were mapped to these categories as other variable definitions did not accurately fit. The variable categorization is shown in Table 15.

Table 15. Variables Included in LSOC Grouped by Categorization

Category	Variable
Autocracy	Polity
Corruption	Govt Effectiveness Corruption
Opportunity	Voice and Accountability
Protest Threat	CPI
Quality of Life	HDI Fertility Rate
Violent Response	Pct Armed Forces Deployed US Troops
Regional Effect	Pct BC
Technology	Internet Users Mobile Cell
Youth Bulge	Youth Labor Youth Bulge Birth Rt

Starting with these 30 variables with the inclusion of time trends, the process of creating a full model included all variables with a significant univariate model test. This resulted in a ten-variable full model. Similar to PSOC, the full model had collinearity amongst many of the variables in the full model. In LSOC, however, the collinearity is reduced by first eliminating all insignificant base variables which corresponded to a trend variable that was also included in the model. This is performed as the momentum of a country's PMESII factors is more evident of its future status than simple, current status. These variables were eliminated one at a time. Remaining insignificant variables were eliminated based on their Wald statistic p-values. Variable elimination was completed one variable at a time, after which the model would be reassessed. Only 3YT CPI and 3YT HDI remained in the reduced model after removing variables with p-values larger than 0.05. The change in parameter estimates for these two variables were 12.4% and 1.4% so none of the other variables in the full model seemed to be confounding. The reduced model also passed the partial likelihood ratio test with a p-value of 0.622 meaning the reduced model's accuracy was comparable to that of the full model within statistical significance. Next, variables that failed the univariate test were tested in the reduced model to see if they were significant, confounding factors. Voice and Accountability was the only significant variable in this test. It was included in the resulting model and the three retained variables' linearity was tested with respect to the logit scaled predicted values. All variables passed the linearity test. No interactions among the three included variables were found significant. The resulting final model for validation is shown in Figure 16.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	26.158912	3	52.31782	<.0001*
Full	14.221788			
Reduced	40.380700			

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	56	14.221788	28.44358
Saturated	59	0.000000	Prob>ChiSq
Fitted	3	14.221788	0.9992

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.94131239	0.8685071	1.17	0.2784
3YT CPI	-0.317652	0.1004542	10.00	0.0016*
3YT HDI	333.179254	112.97871	8.70	0.0032*
Voice and Accountability	0.17801412	0.0730369	5.94	0.0148*

For log odds of 0/1

Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
3YT CPI	1	1	36.6835282	<.0001*
3YT HDI	1	1	38.3802051	<.0001*
Voice and Accountability	1	1	9.90806548	0.0016*

Figure 16. Final Model For LSOC

LSOC should produce a more interpretable model by eliminating the possibility of including noisy or proxy variables. The flow of variables into and out of the model is summarized in Table 16 to demonstrate the benefit of analyzing the inclusion of variables at multiple stages of model building.

Table 16. Flow of Variables into and out of Model in LSOC

Step	Number of Variables
1	30
2a	10
2b	7
2c	3
3a	14
3b	3
4	3
5	3
6	3

Principal Component Regression

Since principal components are linear combinations of all variables, their interpretation requires its own analysis. This section only explains the interpretations of components that passed the univariate tests and remained in a reduced, multivariate model after all insignificant variables were removed. Eleven components passed the univariate test. The full model from these eleven components was reduced by Wald Statistic indication as in PSOC and LSOC. Seven components were retained in the reduced model. To gain an understanding of the meaning captured by the components, variables that were most correlated to the components were analyzed. The variables highly correlated with each component are depicted in Figures 17 through 23 with (+) indicating variables with a positive correlation and (-) indicating negative correlations.

Table 17. Variables Included in PC2

Principal Component 2
Internet Users (-)
Adolescent Fertility Rate (+)
Age Dependency Ratio (+)
Corruption (-)
Government Effectiveness (-)
Mobile Cell (-)
Infant Mortality Rate (+)
Youth Bulge (+)
Working Population (-)
Population Density (-)
Internet Servers (-)
Urban Population (-)

Table 18. Variables Included in PC4

Principal Component 4
Arable Land (+)
Pct Armed Forces (+)
Female Youth Labor Participation (-)
Youth Labor Participation (-)
USAID (+)
Refugee Asylum (+)

Table 19. Variables Included in PC5

Principal Component 5
3YT Death Rate (-)
3YT HDI (+)
3YT Imports (+)
3YT Life Expectancy (+)
3YT Merchandise Imports (+)
3YT Refugee Asylum (+)

Table 20. Variables Included in PC13

Principal Component 13
3YT Arable Land (+)
3YT Computer Exports (+)
3YT CPI (-)
3YT International Tourism (+)
3YT Pct BC (-)
World Press Freedom Index (+)

Table 21. Variables Included in PC17

Principal Component 17
CPI (-)
USAID (+)
3YT Computer Exports (+)
3YT International Tourism (+)
3YT Merchandise Imports (+)
3YT Polity (+)

Table 22. Variables Included in PC21

Principal Component 21
Imports (+)
Youth Labor Participation (+)
Polity (-)
Deployed US Troops (+)
Voice and Accountability (+)
3YT Computer Exoprts (+)
3YT World Press Freedom Index (+)

Table 23. Variables Included in PC29

Principal Component 29
Deployed US Troops (-)
Uneven Development (+)

The proposed understanding of the relationships captured by correlated variables and their effect on the components are summarized in Table 24. This paper cautions against using these understandings as strict definitions for the components, but finds it important to propose possible interpretations for application to our research questions.

Table 24. Definition of PCs

Principal Component	Proposed Definition
2	Increased corruption, Decreasing population
4	Low unemployment, High control of population
5	High quality of life trend
13	Economic growth, Decreased international threat
17	Assistance to struggling countries
21	Openness of government
29	Vacuum of Power

Since there is a reduced understanding of principal component meanings, the reduced model was kept even though there were significant changes in parameter estimates, and it did not look at including components that failed the univariate test. All components included did pass the linearity test with the logit transformation of predicted values. The resulting model is shown in Figure 17. As stated in Section 2.3, some components that explain little of the variance in the PMESII data set (the higher numbered components) were accepted into the final model as they do have a significant relationship to the dependent variable.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	25.886057	7	51.77211	<.0001*
Full	14.494643			
Reduced	40.380700			

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	52	14.494643	28.98929
Saturated	59	0.000000	Prob>ChiSq
Fitted	7	14.494643	0.9959

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	0.88793392	0.5538756	2.57	0.1089
PC2	-0.6144055	0.2125279	8.36	0.0038*
PC4	-1.0792557	0.3595737	9.01	0.0027*
PC5	0.95745083	0.435884	4.82	0.0281*
PC13	1.02355403	0.487972	4.40	0.0359*
PC17	-0.9063549	0.4620497	3.85	0.0498*
PC21	1.30308176	0.6261384	4.33	0.0374*
PC29	2.85894334	1.1452045	6.23	0.0125*

For log odds of 0/1

Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
PC2	1	1	14.7639914	0.0001*
PC4	1	1	23.9959357	<.0001*
PC5	1	1	9.71967249	0.0018*
PC13	1	1	6.23565378	0.0125*
PC17	1	1	5.29779222	0.0214*
PC21	1	1	6.76273967	0.0093*
PC29	1	1	10.7365456	0.0011*

Figure 17. Final Model for PCR

PCR has the potential for improved prediction as all principal components are orthogonal to each other and is another method for capturing factors outside of the scope of our anecdotal hypothesis.

Representative Principal Component Regression

The final method for model building reviewed was initialized following the reduced model of PCR. This method used the correlation definitions of the principal components to select either one or two of the variables highly correlated with the components to represent it as a proxy variable for the component. For the components requiring two variables to define the new indicator, RPCR multiplied these two variables together to use the interaction of these effects as the representative for the components. In effect, the components inform the decisions on which variables to include. RPCR's relationship to PCR is analogous to LSOC as it is a more informed version of the statistically driven PSOC. The variables used as representatives of the components are shown in Table 25.

Table 25. Variables Representing PCs

PC	Representative
2	Corruption
4	Pct Armed Forces * Youth Labor
5	3YT HDI
13	3YT CPI * 3YT Pct BC
17	USAID * 3YT Polity
21	Voice and Accountability * Polity
29	Deployed US Troops * Uneven Development

All of the representatives in Table 25 are included in the multivariate full model, the starting point for RPCR. Representatives were then removed from the model until all parameter estimates had significant Wald statistics (< 0.05). This resulted in two remaining representatives, RPC21 and RPC5. Both representatives passed the linearity test with respect to the logit transformation of predicted values, and the interaction of these two was not found to significantly improve the model. The final resulting model is shown in Figure 18.

Whole Model Test

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	12.732540	2	25.46508	<.0001*
Full	27.648160			
Reduced	40.380700			

Lack Of Fit

Source	DF	-LogLikelihood	ChiSquare
Lack Of Fit	56	26.261866	52.52373
Saturated	58	1.386294	Prob>ChiSq
Fitted	2	27.648160	0.6072

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.9601661	0.4887179	3.86	0.0495*
RPC21	-0.0141702	0.0053691	6.97	0.0083*
RPC5	94.6298769	31.238249	9.18	0.0025*

For log odds of 0/1

Covariance of Estimates

Effect Likelihood Ratio Tests

Source	Nparm	DF	L-R ChiSquare	Prob>ChiSq
RPC21	1	1	10.9141928	0.0010*
RPC5	1	1	19.9285793	<.0001*

Figure 18. Final Model for RPCR

RPCR should provide an interpretable model by using variables from the original data as well as using overarching themes from the conflict dataset as denoted by the PCs.

3.10 Summary

This chapter explains the methods used to analyze and answer the remaining research questions. We begin by describing the open source PMESI data used to fuel the models and provide distinct steps for imputing data in future studies with similar data through the use of multiple imputation by chained equations. The dependent variable is described and its variation from previous studies is defended for the special case of the Arab Spring. To assist the readers' understanding of the models we provide brief descriptions of the variables included in the final models including those calculated from other variables such as Pct BC, trend variables, and PCs. This chapter is concluded with a step-by-step explanation of the four modeling techniques, PSOC, LSOC, PCR, and RPCR.

PSOC provides a purely statistical based methodology that eliminates analyst hypothesis error by including all variables. LSOC allows for research error by only including variables that have previous backing from other studies and political science backing to provide a model with greater causal implications. PCR uses orthogonal PCs as independent variables which should reveal general PMESII themes that effectively predict conflict. RPCR takes base variables that explain the most variation of the PCs in PCR and uses these as independent variables to incorporate the benefits of PCR with a more interpretable model. The following chapter discusses the validation of the resulting models from each of these processes including how these models perform for the use of future prediction in the wake of the Arab Spring.

IV. Results and Analysis

4.1 Overview

This chapter uses the methodology described in Chapter III to answer the remaining research questions through different forms of analysis. This is initiated by running validation techniques on the four models described in the previous chapter. With an understanding of the validity of each model, it is possible to determine which PMESII based hypotheses are supported for conflict proliferation in the Arab Spring. Following this, the analysis provides an interpretation of valid model parameters through multivariate Odds Ratio analysis. Finally, this study runs a primitive predictive analysis on 2016 conflict data and demonstrates how an influence on relevant factors early on in the Arab Spring could have decreased the amount of conflict existing in the region today.

4.2 Model Validation

Four tests were used to determine the validity of the four models, Pearson χ^2 Test, Hosmer-Lemeshow (H-L) Test, Area Under the Curve (AUC) of the Receiver Operator Characteristic (ROC) curve, and classification tables based off of the optimal probability cutoff. The results from the Pearson χ^2 Test, which tests overall model validity, are depicted in Table 26.

Table 26. Pearson χ^2 p-values

Method	Pearson χ^2 p-value
PSOC	0.99968
LSOC	0.99832
PCR	0.01965
RPCR	0.76951

All models except for PCR have high (> 0.7) p-values for this test showing that it is reasonable to suspect that they fit the data in general. With a p-value below 0.05 for PCR however, the test rejects the null hypothesis that the fitted model is correct. The Pearson Test fails when the sum of the squared residuals is large so further analysis investigated the origin of the PCR model's large χ^2 statistic of 75.11 for 52 degrees of freedom. All but one of the squared residuals was less than 5, with the one outlying observation having a squared residual of 58.81. When this observation was removed, the Pearson χ^2 statistic reduced to 16.30 and the p-value improved to 0.999. This observation occurred for the Mauritania 2011 country-year observation when the country experienced conflict yet the model predicted the probability of conflict at 0.016%. Because of the improvement with the exclusion of this observation, model validation statistics are displayed for the PCR model both including and excluding Mauritania 2011 in Table 27.

Table 27. Pearson χ^2 p-values for PCR Outlier Comparison

Outlier Indication	Pearson χ^2 p-value
Full Data	0.01965
Excluding Mauritania 2011	0.99999

The H-L test is similar to the Pearson Test, however, it tests for model fit among multiple bins of estimated probabilities. This study binned the predicted probabilities into ten groups so that each group included six observations. The results of the H-L tests for all four models are displayed in Table 28.

Table 28. Hosmer-Lemeshow Test p-values

Method	Hosmer Lemeshow Test p-value
PSOC	0.99901
LSOC	0.90748
PCR	0.09063
RPCR	0.97997

Once again, all models, except for PCR, indicate good model fit with p-values over 0.8. An investigation of the rejection of the null hypothesis of good model fit for PCR is conducted by observing the C values for each probability group in Table 29.

Table 29. C Values for all Ten Groups Tested in H-L Test

Group	C value
1	0.0027
2	0.0243
3	9.6630
4	0.2355
5	0.7964
6	1.0940
7	1.3822
8	0.2329
9	0.2390
10	0.0053

The third group, which contains the Mauritania 2011 observation, is an outlier among C values at 9.66 alerting us to observe the H-L statistic without the Mauritania 2011 observation. Excluding this observation, the C value for the third group reduces to 0.30 and the p-value of the overall H-L test improves to 0.83. The comparison of the PCR H-L tests with and without Mauritania 2011 is shown in Table 30. The problems caused by this one observation are illustrative of the data granularity limitation. When conducting conflict analysis on a limited region in a limited time span, outliers may have significant leverage effects on model fit and validation.

Table 30. Hosmer-Lemeshow Test p-values for PCR Outlier Comparison

Outlier Indication	Hosmer-Lemeshow Test p-value
Full Data	0.0906
Excluding Mauritania 2011	0.82757

Next, the AUCs from ROC curves were calculated to compare the models' discriminative ability. The results are displayed in Table 31.

Table 31. Area Under the ROC Curve

Method	AUC
PSOC	0.96691
LSOC	0.96190
PCR	0.96181
PCR w/out Outlier	0.9843
RPCR	0.8683

Based on Table 32, LSOC, PSOC, and PCR models have excellent discriminative ability and RPCR has good discriminative ability.

Table 32. AUC Classification Guide

AUC	Classification
0.5	No Discrimination
0.5-0.7	Poor Discrimination
0.7-0.8	Acceptable Discrimination
0.8-0.9	Excellent Discrimination
$\geq$ 0.9	Outstanding Discrimination

Classification tables are closely related to ROC curves. They show sensitivity, specificity, and overall classification accuracy, however they are based on a subjective cutoff to classify observations' probabilities as a positive or negative event. Before running classification tables, the best cutoff value for each model was determined by choosing the cutoff point at the point for which the ROC curve is farthest from the 45° line. This should provide the best balance between sensitivity and specificity and is referred to as optimal. These cutoffs used the coords function from the pROC package in R [50] for calculation. The optimal cutoffs are shown in Table 33.

Table 33. Optimal Probability Cutoffs for Predicted Value Classification

Method	Cutoff Probability
PSOC	0.22785
LSOC	0.63479
PCR	0.46677
RPCR	0.42585

Classification tables were created for all models Using these cutoffs and are shown in Tables 34 through 37, with classification accuracies summarized in Table 38.

Table 34. Classification Table for PSOC

	Predicted 1	Predicted 0
Observed 1	23	1
Observed 0	5	31

Table 35. Classification Table for LSOC

	Predicted 1	Predicted 0
Observed 1	20	4
Observed 0	1	35

Table 36. Classification Table for PCR

	Predicted 1	Predicted 0
Observed 1	22	2
Observed 0	2	34

Table 37. Classification Table for RPCR

	Predicted 1	Predicted 0
Observed 1	19	5
Observed 0	9	27

Table 38. **Classification Accuracies of All Methods**

Method	Classification Accuracy
PSOC	0.90
LSOC	0.9167
PCR	0.933
RPCR	0.767

The classification tables not only show the accuracy of the four models, but also show the ways in which the model fails to predict conflict. Through analysis of the false-positive and false-negative rates of each model, one can better understand the weaknesses of the four models. One conservative standpoint of conflict prediction might prefer a model with a larger proportion of false-positives to false-negatives in order to avoid missing a high-intensity, violent conflict. Only the LSOC model failed to achieve this conservative approach by overlooking two years of conflict in Tunisia and one in Syria and Mauritania. Future research should investigate what may have caused the discrepancy between Tunisia's improving CPI and HDI trends as well as its Voice and Accountability and its high conflict rating.

Another approach to conflict prediction might prefer a higher proportion of false-negatives to false-positives due to the reality of budget and manpower scarcity. If this conflict prediction tool is to be used to inform decision makers on policy goals and vectors in future environments of abnormal conflict transitions, it may be more useful to highlight only the most serious cases of conflict. The PSOC and RPCR models have high rates of false-positives, incorrectly specifying a year in conflict 8.3% and 15% of the time respectively. This might degrade the reliability of these models in the eyes of the decision maker. Although RPCR's false-positives occur over a varied number of countries and years, three of the five false-positives in the PSOC model occurred for Jordan and another occurred in Kuwait, both countries that saw no years of high-intensity, violent conflict in the five-year time span. Future research should investigate these failures for the otherwise accurate PSOC model.

Taking into account both accuracy and validation analysis, PCR is removed from consideration as the accepted model for analysis. Although this model achieves the greatest accuracy as seen in Table 38, PCR cannot be validated due to its poor prediction of the Mauritania 2011 observation. In many circumstances, this observation is seen as a true outlier in the data and could be thrown out to more accurately model the situation. In this data however, Mauritania 2011 represents the first year of protests in the country, starting in January 2011, and is not seen as an outlier in any of the three other models. This was caused by PCR's inability to recognize the magnitude of protesters' demands as described in Section 2.2 as there were two suicide bombers and some conflict between citizens and police caused by protests which led to a HIIK conflict classification of four [15].

RPCR was removed from consideration as it achieved around 15% worse classification accuracy than that of PSOC and LSOC. These remaining two models, sharing two of their three variables, achieved similar classification accuracies, and were both validated in all tests. The LSOC model, however, was chosen as the accepted model for its improved, albeit comparable, accuracy and increased interpretability. The LSOC bases the candidate variable pool on the hypothesis outlined in Section 2.2 which relies on expert research and factual events of the Arab Spring. The validation of this model through logistic regression proves that PMESII factors can demonstrate a reliable analysis of conflict in the Arab Spring. LSOC is the suggested answer to research question S2: what is the most effective method for model building to capture the tendency of a country to fall into conflict?

4.3 Interpretation of Important Variables

Using the principles of logistic regression, this study can interpret the effects that each factor has on the model's probability prediction. The coefficient estimates in logistic regression cannot be translated to their real-life effects as simply as linear regression due to its nonlinear nature. In linear regression, coefficient estimates are interpreted as the effect that a one unit increase in a factor has on the dependent variable. In logistic regression, however, one must exponentiate the scaled coefficient estimates of continuous variables to determine their effect on the probability of the dependent variable [23]. This exponentiated figure is called the Odds Ratio and demonstrates how a given increase in the continuous independent variable affects the odds of the dependent variable being a success or, in our context, in conflict. Odds ratios can be better understood following the direction in Equation 18.

$$[H]\hat{\beta}: \begin{cases} \hat{\beta} \text{ decreases odds of } \hat{Y}, & \text{if } \widehat{OR} < 1 \\ \hat{\beta} \text{ has no effect on odds of } \hat{Y}, & \widehat{OR} = 1 \\ \hat{\beta} \text{ increases odds of } \hat{Y}, & \widehat{OR} > 1 \end{cases} \tag{18}$$

Odds Ratios for continuous variables explain how a given increase, c, in the independent variable affects the odds of the dependent variable. This value c, should be individually specified for each independent variable to reflect a reasonable change in that variable. A reasonable change should not be larger than the variable's range, but should also be large enough to understand how an actionable change might affect the probability prediction. The c values for each variable in the four final models are displayed along with each variable's range in Table 39.

Table 39. c-values for each Variable Used in Final Models

Variable	Range	c
3YT CPI	74.59	1
3YT HDI	0.135	0.01
3YT Refugee Origin (per million)	9.742	0.1
Voice & Accountability	74.99	1
RPCR 5	0.135	0.1
RPCR 21	726.1	1
PC 2	13.07	0.1
PC 4	10.30	0.1
PC 5	10.29	0.1
PC 13	6.572	0.1
PC 17	5.820	0.1
PC 21	5.138	0.1
PC 29	3.4	0.1

The c values and coefficient estimates are used to find the Odds Ratios according to Equation 19.

$$\widehat{OR} = e^{c\hat{\beta}_1} \tag{19}$$

An important attribute of Odds Ratios when analyzing the effects of variables in models is their limitation in the multivariable case. The Odds Ratios in Equation 19 assume constant values of all other variables in the model. By testing the change in coefficient estimates from the univariate to multivariate models it is possible to determine if a variable is a confounding effect for other variables in the model or if an interaction of variables in the model is necessary. Although all of the modeling processes sought to account for confounders by adding variables back in after the reduced model, testing the change in coefficient estimates allows the analyst to observe which variables were confounding and which were the main effects. For PSOC and LSOC, 3YT CPI and 3YT HDI are the main effects of both models, 3YT Refugee Origin (per 1 million) is a confounding factor for both in PSOC, and Voice and

Accountability is a confounding factor for both in LSOC. RPCR21 is a confounding factor of main effect RPCR5. All PCs in the PCR model are main effects.

By accepting the LSOC model as the most reliable and interpretable model, it is possible to analyze the important measures for explaining conflict. The Odds Ratios for this model indicate that the odds of conflict in a given country-year observation are increased with an increasing trend in the Consumer Price Index, a decreasing trend in the Human Development Index, both supplemented by a decreased Voice and Accountability rating. On a generalized level, this appears to indicate that a decreasing quality of life and a low ability to affect change in a country increases the probability of conflict. This finding provides support for two of the three factors for conflict in the Anecdotal Hypothesis in Section 2.2. It supports the idea that Arab Spring conflict is more likely when factors affecting the population, which can be thought of as quality of life factors, are decreased over the three-year trends in our model. It also agrees that a decreased willingness of a government to concede to protesters, manifested by an inverse relationship to Voice and Accountability, leads to violent clashes with the population which undermine the government and spark the violence for war. This answers research question P2: what open source PMESII factors affect the selected Arab Spring nations' tendencies to transition into and out of conflict?

4.4 Answering Research Questions

The lack of validation data limits the analysis of the modeling. This is a necessary setback, however, due to the limited real data of Arab Spring country-year observations. A primitive prediction of 2016 uses the predicted probabilities calculated from the 2015 independent variables to provide pseudo test set validation. While it is possible that the independent variables could have changed drastically over the one-year period, the absence of HDI 2016 data prohibited accurate predictions for 2016. The results of the primitive prediction from LSOC are displayed in Table 40. The high primitive prediction classification accuracy (91.7%) of the LSOC model addresses research question P3: Can Arab Spring nations be grouped into two groups: at-risk for escalated conflict and not at-risk for escalated conflict?

Table 40. LSOC 2016 Primitive Prediction Classification Table

	Predicted 1	Predicted 0
Observed 1	5	1
Observed 0	0	6

The final addition of analysis shows how much each relevant variable would have been needed to be adjusted to reduce all countries out of high-intensity, violent conflict. Using the LSOC optimal cutoff of 0.63479, Table 41 shows the increase or decrease needed for each variable to force all countries out of conflict for 2011 holding all other factors constant.

Table 41. Change in Variables Needed to Predict No Countries In Conflict in 2011

Country	3YT CPI	3YT HDI	Voice & Accountability
Egypt	-10.3	0.01	18.3
Jordan	-3.6	0.0034	6.4
Libya	-51.5	0.05	91.8
Syria	-0.7	0.0007	1.2
Yemen	-0.6	0.0006	1.1

These findings address research question P4: could the probability of long-term, escalated conflict have been decreased by altering certain PMESII factors? The method used to evaluate these changes in relevant variables to decrease conflict prediction provides a way for decision makers to make quantitative goals in similar, future environments. In the world of limited resources that the US military faces, it is also useful to know which countries have the most dire needs for change and which countries could be swayed from conflict with the smallest magnitude of change.

4.5 Summary

This chapter applies statistical tests to the models obtained from the previous chapter and uses both statistical and analytic insight to relate it to the research questions and hypothesis. We validate all four models using the Pearson χ^2 Test and Hosmer-Lemeshow Test, and compare accuracies using Classification Tables and AUCs from ROC curves. The data generated from these tests proved LSOC to be the best tested modeling technique, and allowed adoption of the three variable model resulting from LSOC as the proposed model for further analysis. The elements of the LSOC model are related to the anecdotal hypothesis and show how it can be used for prediction and analysis in future environments of increased, regional conflict outbreak.

V. Conclusions

5.1 Overview

This chapter organizes the findings of the research efforts. It provides evidence related to the two statistical and four political research questions and how they relate to the overall problem statement. This leads into discussion on the practical uses and implications of the research in the US military. This chapter concludes with useful avenues for furthering this research in the future.

5.2 Insights from Research Questions

In concluding this thesis, this section first addresses the statistical and political research questions to summarize our results.

Statistical Research Questions

S1: What is the best method for imputing missing PMESII data, even when a large portion of the most recent data is missing?

Multiple imputation by chained equations is shown to be a useful method of imputation for a dataset with many variables in Section 3.5. The proposed method of imputing based off of a parametric distribution formed by the non-missing data for each country is especially proficient at imputing data when an entire year is missing from a variable. This study does not defend the application of imputation techniques to variables in which any country had no data for all years studied which limited the pool of candidate variables.

S2: What is the most effective method for model building to capture the tendency of a country to fall into conflict?

Based on all validation statistics in Section 4.2 PCR is the most accurate model, however its problems with the Mauritania 2011 observation and the lack of clarity in the principal components makes this model harder to accept for practical use. The next most accurate model is LSOC which attained 92% training set accuracy, and passed all validation tests. The incorporation of only hypothesized factors in the LSOC model also lends this model to interpretability and causal analysis. The LSOC method is considered to be the best tested method for conflict modeling in the Arab Spring context.

Political Research Questions

P1: How can nations be identified as being affected by the Arab Spring?

There is precedence in answering this question, found during the Literature Review efforts. Costello et al. [6] used press reports to determine the number of violent and non-violent protests in the Middle East and North African region before and after the onset of the Arab Spring in 2011. They hold that the nations experiencing an uptick in either violent or non-violent protests corresponded to those nations experiencing other characteristics deemed unique to the Arab Spring such as a call for reform and increased human rights. This study agrees with Costello et al. [6] and incorporates all countries included in what they call the "Arab Awakening" as well as Saudi Arabia, which experienced a brief uptick in protests from the Arab Spring movement in 2011 as the countries affected by the Arab Spring. The defining characterization of these countries is the experience of increased protests motivated by a similar, regional cause.

P2: What open source PMESII factors affect the selected Arab Spring nations' tendencies to transition into and out of conflict?

The dataset used for logistic regression included 53 independent variables spanning all aspects of the PMESII spectrum. The most relevant factors in determining Arab

Spring nations' tendencies into and out of conflict should therefore be captured in the most successful logistic regression model. By accepting the LSOC model, this study proposes both raw and hypothesized factors that should best capture these tendencies. The variables in the LSOC model, which was the best mix of accuracy and interpretability, were a three-year trend of the Human Development Index, a three-year trend of Consumer Price Index, and the World Governance Indicator, Voice and Accountability. These three variables capture two main ideas that tend countries towards conflict, decreasing quality of life, and a lack of legal political opportunity for citizens to voice their complaints.

P3: Can nations receiving a conflict shock be grouped into two groups: at-risk for escalated conflict and not at-risk for long-term, escalated conflict over a five-year period following the shock?

Shallcross [4] used PSOC as a model building technique for a conflict-prediction, logistic regression model. We test this methodology along with three other model building strategies not found in the conflict-prediction literature to find the best model building strategy to classify countries into either conflict or non-conflict predicted states. LSOC, a technique closely related to PSOC with a more hypothesis-based approach, is found to have the best combination of accuracy and interpretability. The LSOC model correctly classifies 92% of training set observations. There is also evidence through primitive predictive measures that the LSOC model can predict future conflict in the Arab Spring with an accuracy of 92%. There is no current data to perform predictions of further conflicts and this study doubts the usefulness of such predictions as the Arab Spring environment may have subsided into expected conflict behavior.

P4: Could the probability of long-term, escalated conflict have been decreased following the Arab Spring by altering certain PMESII factors?

By applying the Odds Ratio values for each variable in the LSOC model, it is possible to recalculate the predicted probabilities for individual country-year observations with altered variable levels. The efficacy of affecting change to the conflict environment is summarized in Table 41. The results indicate that some countries, like Syria and Yemen, could have been realistically affected with foreign input to change their conflict status. Others, like Libya, would require more unrealistic goals for improvement to remove them from being at risk for high-intensity, violent conflict. It must also be noted that the inclusion of three-year trend variables in the model indicates that resolution of conflict in the Arab Spring takes time.

5.3 Summary of Research Efforts

After data cleaning and innovative imputation of PMESII factors, this research tested 53 independent variables to obtain a useful model for Arab Spring conflict. Three modeling techniques, unused in previous conflict research, were tested alongside PSOC, and restrict the data to Arab Spring country-year observations to fuel logistic regression analysis which has failed to achieve conflict prediction accuracies of 80% in this historical scenario in previous studies. By focusing on countries that were, in fact, affected by the Arab Spring separately, and by expanding the candidate independent variable set from previous studies, this study obtains a logistic regression model with training set classification accuracies in excess of 90%. The methods used to obtain the data and models, along with the success of our best model allow this study to address our initial six research questions.

5.4 Future Research

This study is one of the first logistic regression studies of the conflict behavior of the Arab Spring after its onset. As an early analysis of this topic, future studies should examine our limitations in data and scope.

Data Expansion

There were two types of limitations caused by our data availability. The first was due to a lack of independent variables or missing data for a whole country in present variables that would have better addressed the hypothesis. Of note, this study did not have access to variables that represented the strength of the mechanisms that clashed with protesters such as police strength or those that accounted for the prevalence of terrorism in the country. In the discussion on imputation, it was not assumed defendable to impute data for a country that had no historical records for a given variable, however this assumption could be investigated to help conflict studies incorporate a broader range of PMESII factors. This might be accomplished by finding regional relationships between similar variables to those that are missing a full country's data.

The second data limitation that would provide a meaningful contribution to Arab Spring analysis, would be to observe conflict on a more granular scale. By analyzing how conflict reacts to PMESII factors on a sub-country or monthly scale instead of a country-year scale it might be possible to determine more specifically what themes caused certain parts of the Arab Spring region to further destabilize and others to settle into lower levels of conflict. A more granular analysis would also increase the sample size of the study which would allow for a meaningful training set analysis to further test models.

Scope Expansion

There are three areas suggested to broaden the scope of this research. The first is determining how to best incorporate a separate methodology, such as this Arab

Spring model, to existing conflict prediction models by identifying when a unique conflict climate arises and adjusting conflict probabilities accordingly. Second, is the possibility of melding this study's methodology during a setting of high conflict transition with that of Shallcross [4] which applies conflict transition probabilities to a Markov Chain model to analyze the flow of countries into and out of conflict over multiple years. Finally, beyond international conflict prediction, this study shows promise in application to smaller scale destabilization from initial protests. In recent history protests in Ferguson, Baltimore, and Charlottesville have led to significant property damages and casualties. A similar methodology as that of this paper could be applied to this setting to understand how best to respond to future protests in the US.

5.5 Significance of Research

This study provides significant insight to both the statistical field of conflict prediction and to the political science field in the understanding of the Arab Spring. Implementation of the MICE imputation methodology using given-data parametric distributions should provide realistic results for time-series, conflict data. Following this, model building testing shows that although high training set accuracy can be achieved using pure statistical methods, a hypothesis-backed model may provide similar accuracy measures and allows for greater understanding of conflict. On the political side, this research reveals the contributing factors to the behavior of Arab Spring countries' conflict tendencies. It is possible that future regional conflict shocks will be focused on factors other than quality of life and government repression, which analysts should address by using this model building methodology. The factors included in models from such a methodology create more concrete goals for decision makers to work towards for similar situations in the future. Along with understanding the pertinent factors to affect conflict in unusual conflict climates, this research suggests that through the use of logistic regression it is possible to determine which countries can be realistically affected to decrease their conflict threat levels. With a grasp on which countries are at risk and which countries can be most easily adjusted to take them out of the at-risk classification, decision makers can make better-informed decisions on where and in what way to supply resources to the region.

Appendix A. Variables Considered in Modeling

Table 42. AUC Classification Guide

Variables
Adolescent fertility rate (births per 1,000 women ages 15-19)
Age dependency ratio (% of working-age population)
Arable land (% of land area)
Armed forces personnel (% of total labor force)
Birth rate, crude (per 1,000 people)
Computer, communications and other services (% of commercial service exports)
Consumer price index (2010 = 100)
Corruption Perceptions Index (CPI)
Death rate, crude (per 1,000 people)
Deployed US Troops
Fertility rate, total (births per woman)
Fragile States Index (FSI): Uneven Development
Freedom in the World: Civil Liberties
Government Effectiveness: Percentile Rank
Human Development Index (HDI)
Import volume index (2000 = 100)
Inflation, consumer prices (annual %)
International tourism, expenditures (current US$)
Internet users (per 100 people)
Labor force participation rate for ages 15-24, female (%) (modeled ILO estimate)
Labor force participation rate for ages 15-24, total (%) (modeled ILO estimate)
Life expectancy at birth, total (years)

Merchandise imports (current US$)
Mobile cellular subscriptions (per 100 people)
Mortality rate, infant (per 1,000 live births)
Percent Border Conflict
Population ages 0-14 (% of total)
Population ages 15-64 (% of total)
Population ages 65 and above (% of total)
Population density (people per sq. km of land area)
Population growth (annual %)
Refugee population by country or territory of asylum
Refugee population by country or territory of origin
Revised Combined Polity Score for Time-Series Analysis
Secure Internet servers (per 1 million people)
Time required to enforce a contract (days)
Time required to start a business (days)
Time to prepare and pay taxes (hours)
Urban population (% of total)
USAID Economic Assistance ($US)
Voice and Accountability: Percentile Rank
World Press Freedom Index (PFI) Standardized Score

Bibliography

1. "Conflict barometer 2016," tech. rep., Heidelberg Institute for International Conflict Research, 2016.
2. F. G. G. III, "Why middle east studies missed the arab spring: The myth of authoritarian stability," *Foreign Affairs*, vol. 90, pp. 81–84, 2011.
3. B. C. Boekestein, "A predictive logistic regression model of world conflict using open source data," Master's thesis, Air Force Institute of Technology, 2015.
4. N. J. Shallcross, "A logistic regression and markov model for the prediction of nation-state violent conflicts and transitions," Master's thesis, Air Force Institute of Technology, 2016.
5. B. D. Leiby, "A conditional logistic regression predictive model of world conflict considering neighboring conflict and environmental security," Master's thesis, Air Force Institute of Technology, 2017.
6. M. J. Costello, C. Jenkins, and H. Aly, "Bread, justice, or opportunity? the determinants of the arab awakening protests," *World Development*, vol. 67, pp. 90–100, 2015.
7. M. Jamshidi, *The Future of the Arab Spring.* Butterworth-Heinemann, 2013.
8. K. Weyland, "The arab spring: Why the surprising similarities with the revolutionary wave of 1848?," *Perspectives on Politics*, vol. 10, no. 4, pp. 917–934, 2012.
9. S. A. Tobin, "Jordan's arab spring: The middle class and anti-revolution," *Middle East Policy*, vol. 19, no. 1, pp. 96–109, 2012.
10. A. Engel, "The islamic state's expansion in libya," tech. rep., The Washington Institute for Near East Policy, 2015.
11. K. McEvers, "Bahrain: The revolution that wasn't : Npr," *National Public Radio*, 2012.
12. G. Steinberg, "Leading the counter-revolution. saudi arabia and the arab spring," tech. rep., German Institute for International and Security Affairs, 2014.
13. H. Albloshi, "Sectarianism and the arab spring: The case of the kuwaiti shi'a.: Discovery service for air force institute of technology," *Muslim World*, vol. 106, pp. 109–126, 2016.
14. S. I. Bergh and D. Rossi-Doria, "Plus ça change? observing the dynamics of morocco's 'arab spring' in the high atlas," *Mediterranean Politics*, vol. 20, no. 2, pp. 198–216, 2015.

15. C. Toulmin, "Conflict analysis of mauritania expert contributors," tech. rep., Governance and Social Development Resource Centre (GSDRC), 2014.

16. M. Valeri, "Simmering unrest and succession challenges in oman january," tech. rep., Carnegie Endowment for International Peace, 2011.

17. "Syria profile - timeline - bbc news," *British Broadcasting Corporation*, 2017.

18. R. Popp, "War in yemen: Revolution and saudi intervention," *Center for Security Studies Analyses in Security Policy*, no. 175, pp. 1–4, 2015.

19. H. Urdal, "A clash of generations? youth bulges and political violence," *International Studies Quarterly*, vol. 50, pp. 607–629, 2006.

20. G. Wolfsfeld, E. Segev, and T. Sheafer, "Social media and the arab spring," *The International Journal of Press/Politics*, vol. 18, no. 2, pp. 115–137, 2013.

21. H. H. Khondker, "New media, political mobilization, and the arab spring," *International Encyclopedia of the Social & Behavioral Sciences*, pp. 798–804, 2015.

22. R. Rifai, "Timeline: Tunisia's uprising," *Al Jazeera*, 2011.

23. D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*. Hoboken NJ: John Wiley and Sons, Inc., third ed., 2013.

24. "Assumptions of logistic regression." `https://http://www.statisticssolutions.com/wp-content/uploads/wp-post-to-pdf-enhanced-cache/1/assumptions-of-logistic-regression.pdf`.

25. "The fundamental difference between principal component analysis and factor analysis." `https://www.theanalysisfactor.com/`. Accessed: 2017-7-25.

26. A. S. Had and R. F. Ling, "Some cautionary notes on the use of principal components regression," *The American Statistician*, vol. 52, no. 1, pp. 15–19, 1998.

27. R. Shearer and B. Marvin, "Recognizing patterns of nation-state instability that lead to conflict," *Military Operations Research*, vol. 15, no. 3, pp. 17–30, 2010.

28. S. Rao, "Conflict analysis of mauritania," tech. rep., Governance and Social Development Resource Centre (GSDRC), 2014.

29. "Joint intelligence preparation of the operational environment," tech. rep., US DOD Joint Chiefs of Staff, 2014.

30. H. Kang, "The prevention and handling of the missing data," *Korean journal of anesthesiology*, vol. 64, no. 5, pp. 402–406, 2013.

31. S. van Buuren and K. Groothuis-Oudshoorn, "mice: Multivariate imputation by chained equations in r," *Journal of Statistical Software*, vol. 45, no. 3, pp. 1–67, 2011.

32. S. van Buuren, "Multiple imputation of discrete and continuous data by fully conditional specification," *Statistical Methods in Medical Research*, vol. 16, pp. 219–242a, 2007.

33. J. L. Schafer, "Multiple imputation: a primer," *Statistical Methods in Medical Research*, vol. 8, no. 1, pp. 3–15, 1999.

34. L. Doove, S. V. Buuren, and E. Dusseldorp, "Recursive partitioning for missing data imputation in the presence of interaction effects," *Computational Statistics and Data Analysis*, vol. 72, pp. 92–104, 2014.

35. G. Vink, L. E. Frank, J. Pannekoek, and S. van Buuren, "Predictive mean matching imputation of semicontinuous variables," *Statistica Neerlandica*, vol. 68, no. 1, pp. 61–90, 2014.

36. K. Abayomi, A. Gelman, and M. Levy, "Diagnostics for multivariate imputations," *Journal of the Royal Statistical Society*, vol. 57, no. 3, pp. 273–291, 2008.

37. S. Engmann and D. Cousineau, "Comparing distributions: The two-sample anderson-darling test as an alternative to the kolmogorov-smirnoff test," *Journal of Applied Quantitative Methods*, vol. 6, no. 3, pp. 1–17, 2011.

38. "Syria inflation rate." `https://tradingeconomics.com/syria/inflation-cpi`, 2017. Accessed: 2017-10-12.

39. "World bank open data." `https://data.worldbank.org/`, 2017. Accessed: 2017-6-12.

40. "Consumer price index manual : Theory and practice." `https://www.imf.org/en/Publications/Manuals-Guides/Issues/2016/12/30/Consumer-Price-Index-Manual-Theory-and-Practice-17165`, 2017. Accessed: 2017-11-15.

41. "Human development index." `http://hdr.undp.org/en/content/human-development-index-hdi`, 2017. Accessed: 2017-11-15.

42. "World governance indicators." `http://info.worldbank.org/governance/wgi/#home`, 2017. Accessed: 2017-9-28.

43. "The polity project." `http://www.systemicpeace.org/polityproject.html`, 2017. Accessed: 2017-11-15.

44. "Corruption perceptions index." `https://www.transparency.org/news/feature/corruption_perceptions_index_2016`, 2017. Accessed: 2017-11-15.

45. “Armed forces personnel (https://tradingeconomics.com/syria/inflation-cpi, 2017. Accessed: 2017-11-17.

46. “Labour force participation rate.” http://www.ilo.org/ilostat/faces/oracle/webcenter/portalapp/pagehierarchy/Page3.jspx?MBI_ID=15&_afrLoop=200845983479934&_afrWindowMode=0&_afrWindowId=z2w6itc87_1#!%40%40%3F_afrWindowId%3Dz2w6itc87_1%26_afrLoop%3D200845983479934%26MBI_ID%3D15%26_afrWindowM, 2017. Accessed: 2017-11-17.

47. “UNHCR - who we help.” http://www.unhcr.org/en-us/who-we-help.html, 2017. Accessed: 2017-11-17.

48. “USAID.” https://www.usaid.gov/, 2017. Accessed: 2017-11-17.

49. “Fragile states index.” http://fundforpeace.org/fsi/, 2017. Accessed: 2017-11-15.

50. X. Robin, N. Turck, A. Hainard, N. Tibereti, F. Lisacek, J.-C. Sanchez, and M. Muller, “proc: an open-source package for r and s+ to analyze and compare roc curves,” *BMC Bioinformatics*, vol. 12, p. 77, 2011.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704–0188

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
22–03–2018	Master's Thesis	Aug 2016 — Mar 2018

4. TITLE AND SUBTITLE

LOOKING PAST THE SPARK TO FIND THE FUEL OF THE ARAB SPRING FIRE

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Luke M. Brantley, 2d Lt, U.S. Air Force

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT-ENS-MS-18-M-111

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Intentionally Left Blank

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A:
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The field of statistical conflict prediction addresses region-wide analysis in eras of stable conflict and peace. This study improves upon those prediction rates in times of volatile conflict and peace seen during the Arab Spring of 2011 to 2015. During this time, higher rates of conflict transition in certain Middle Eastern and North African countries occurred than normally observed in previous studies. Due to the fact that previous prediction models decrease in accuracy during times of volatile conflict transition and since the proper strategy for handling the Arab Spring has been highly debated, this study considers alterations to previous studies to understand the effects of the Arab Spring on conflict prediction over a five-year period. This study identifies which countries were affected by the Arab Spring, and then apply logistic regression to predict a country's tendency to suffer from high-intensity, violent conflict. A large number of open-source variables are incorporated by implementing an imputation methodology useful to conflict prediction studies in the future. The imputed variables are implemented in four model building techniques: Purposeful Selection of Covariates, Logical Selection of Covariates, Principal Component Regression, and Representative Principal Component Regression resulting in accuracies exceeding 90%. Analysis of the models produced by the four techniques supports hypotheses which propose political opportunity and quality of life factors as causations for increased instability following the Arab Spring.

15. SUBJECT TERMS

Logistic Regression, Assessment, Arab Spring, Imputation

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Dr. Darryl K. Ahner, AFIT/ENS
					19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	107	(937)823-9278; Darryl.Ahner@afit.edu